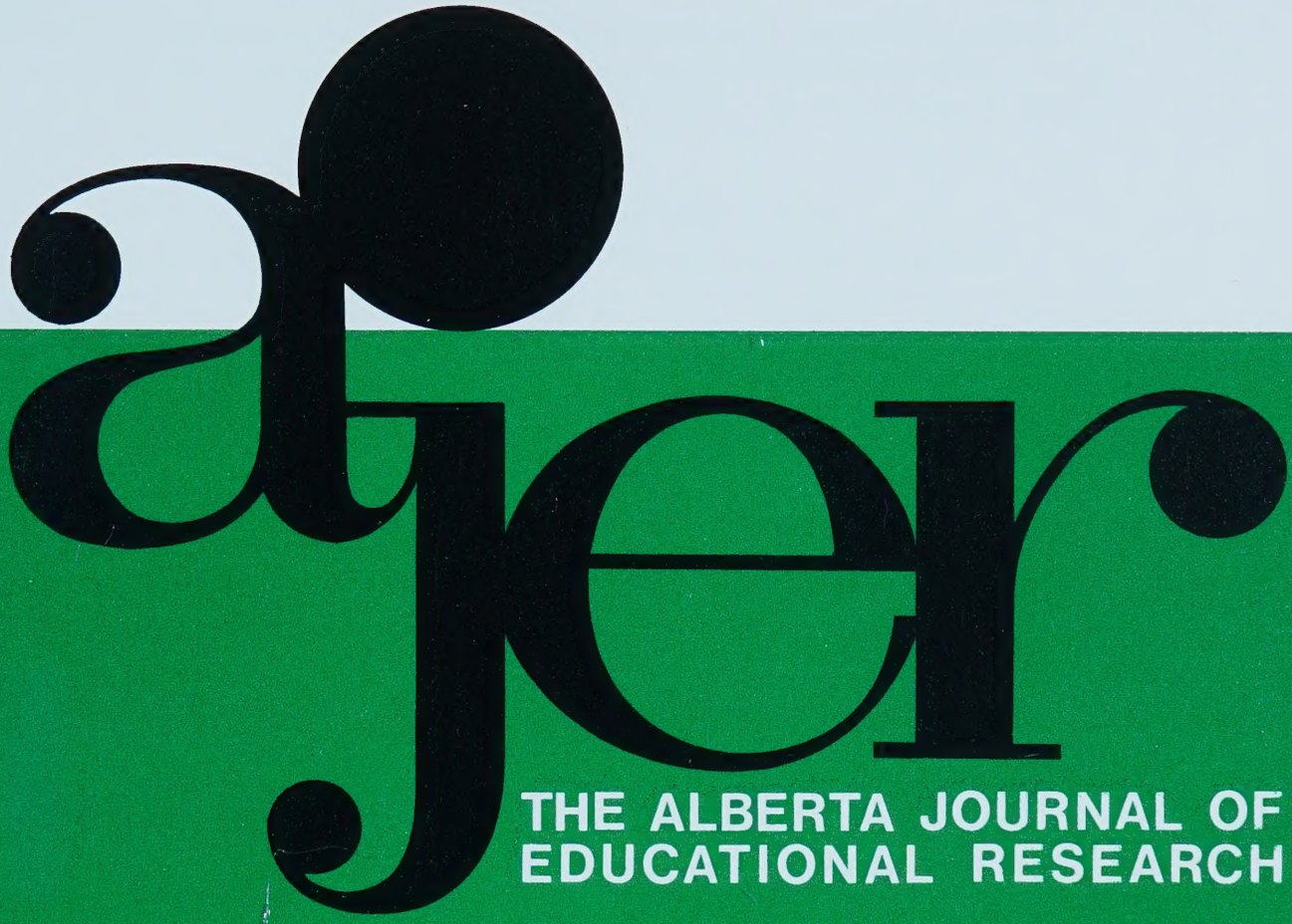

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXIV NUMBER 2 JUNE 1978

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

::::::::::::

Published quarterly in March, June, September and December
by the
Faculty of Education
of
The University of Alberta

::::::::::::

Faculty Publications Committee

M. A. Assheton-Smith	*W. T. Fagan*	*R. G. Martin*
N. C. Bhattacharya	*H. W. Hodysh*	*E. Miklos*
J. Chambers	*W. W. Laing*	*C. H. Preitz*

Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

Chairman: *E. Miklos* Editor: *Henry W. Hodysh*
Secretary: Mrs. A. Onishenko

::::::::::::

AJER gratefully acknowledges support from Canada Council and AACES.
AJER is indexed in the *Canadian Education Index;* appropriate articles are abstracted in *Educational Administration Abstracts, Psychological Abstracts,* and *Sociology of Education Abstracts.*

The subscription rate is $8.00 per year; single copies are $2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research.* All back issues are available in bound and unbound volumes; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications to the Editor, *The Alberta Journal of Educational Research,* Faculty of Education, The University of Alberta, Edmonton, Alberta, Canada T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436

ISSN 0002-4805

The Alberta Journal of Educational Research

Volume XXIV, No. 2 *June, 1978*

CONTENTS

Convergent and Discriminant Validity of Classroom Observation Instruments: Conceptual Background, Critique, and a Case in Point 69
H. Munby and R. Wilson

Anxiety Reduction Correlates of Adult Learners: A Longitudinal Study 81
Y. L. J. Lam

Skills Underlying Coloured Progressive Matrices 94
J. R. Kirby and J. P. Das

Cross-modal and Intramodal Matching: Relationships to Simultaneous and Successive Syntheses and Levels of Performance Among Three Intelligence Groups 100
R. F. Jarman

The Effect of Sentence-Combining Practice on Written Compositions and Reading Comprehension 113
W. D. Smart and L. O. Ollila

The Relationship Between Teachers' Experimental and Dogmatic Attitudes, and their Pupils' Self-Concept 121
G. C. S. Cheong and E. P. Wadden

BOOK REVIEWS

New Career Options for Women: A Counselors' Sourcebook, Helen S. Farmer and T. E. Backer 126
Reviewed by R. A. Sydie

Adult Education as a Field of Study and Practice, D. Campbell 128
Reviewed by T. C. Byrne

People, Culture and Learning, J. W. Friesen 130
Reviewed by J. R. Young

Cognition, Curriculum, and Comprehension, J. T. Guthrie (Ed.) 131
Reviewed by W. T. Fagan

List of Books Received 135

FACULTY OF EDUCATION
The University of Alberta

Announcement

of the

CENTRE FOR RESEARCH IN TEACHING

The *Centre for Research in Teaching* has been established in the Faculty of Education of the University of Alberta. The Centre, which is the first of its kind in Canada, will have linkages with at least two similar units in the United States. Dr. D. A. MacKay has been appointed Coordinator of the Centre.

The Centre has as its main purposes, for the period July, 1977, to August, 1980, the following:

1. To facilitate involvement of staff members of the Faculty of Education, as well as other scholars, in research in teaching and teacher education.
2. To disseminate research findings that have particular significance for teachers, teacher educators, and research scholars.
3. To plan and implement a program of coordinated research in teaching and teacher education that will involve staff and graduate students of the Faculty of Education.
4. To ensure continuation and further expansion of this line of research beyond the three-year "seed period" by developing and presenting proposals for major funding to various external agencies.

Anyone wishing to contact the Centre may write to the following address:

Dr. D. A. MacKay, Coordinator
Centre for Research in Teaching
Room 234, Education I
The University of Alberta
Edmonton, Alberta, Canada
T6G 2G5

The Alberta Journal of Educational Research *Vol. XXIV, No. 2, June, 1978*

HUGH MUNBY
and
ROBERT WILSON
Queen's University

Convergent and Discriminant Validity of Classroom Observation Instruments: Conceptual Background, Critique, and a Case in Point

An epistemological interpretation of the Campbell and Fiske model is offered to show its usefulness in validating classroom observation instruments. This interpretation is used to critique a recent validation study, and is then applied to the problem of validating the construct "View of Science," using a classroom observation instrument.

The validation study consisted of submitting a composite of three written test instruments to 47 science teachers, and analyzing their teaching behaviour using three corresponding observation instruments. Eighteen of the 21 possible correlations in the Campbell and Fiske matrix achieved appropriate levels of significance. This result demonstrates the plausibility of using the multitrait-multimethod model for validating observation instruments, and establishes a preliminary validity for the construct "View of Science." (Dr. Munby is Associate Professor of Curriculum at the Faculty of Education at Queen's University; Dr. Wilson is Associate Dean of Research, Development, and Continuing Education in the same faculty.)

This investigation is concerned with the validity of classroom observation instruments, and with the theoretical significance and practical application of the multitrait-multimethod validation model. Questions about the validity of all measures in educational research persist for at least two reasons. First, it is evident that the link between measured behaviours and our conceptualization of what these show or can be taken to mean is a tenuous one, at best, and this is a feature of the social sciences which renders its epistemology more problematic than that of the physical sciences. Second, and related to the first, in many areas of educational research there is still inadequate heed taken of the need for validity of instruments. For example, the prevalence of author-constructed instruments of unknown characteristics which are employed to tap constructs having unknown conceptual relationships to

familiar constructs is a frequently cited complaint in reviews of science education research (Trowbridge, 1972; Novak, 1973). Likewise, while there tends to be much information available about the *reliability* of classroom observation instruments, there appears to have been substantially less attention given to the *validity* of these devices (Borich & Malitz, 1975). It is as if we can agree to term a specific teaching behaviour "lecturing," but still not be sure that the construct accurately captures the act.

It is within this context of uncertainty about validity that the study reported here is lodged, not with a view to extinguishing the uncertainty but to arguing theoretically and instancing empirically that it can begin to be addressed. For it will be shown here that a theoretically derived construct, measured by a classroom observation scheme, can be validated against observation and written measures so that its theoretical relationships with other constructs are documented and its characteristics in use are delineated. The villain of this piece is the construct "View of Science," described below; it is the purpose of this study to validate this construct using the Campbell and Fiske multitrait-multimethod matrix. Before the design and results of this study can be discussed it is necessary to explore the elementary epistemology that informs the concept of validity in general and the Campbell and Fiske matrix in particular. This is necessary for distinguishing a conceptually inappropriate use of the model (in a study critiqued later) from a conceptually appropriate use (as demonstrated in the validity study reported here).

A Basic Epistemology for Construct Validation

The standard account of validity centres on the general notion that the validity of a measuring instrument represents the correspondence between what in fact the numbers are measures of and what they are thought to be measures of. A major implication of this definition is that it makes no sense to consider that the validity of an instrument is conceptually distinguishable from the validity of a construct. That is to say, there is no way of ever knowing for certain what, in fact, numbers generated by social science instruments *are* measures of. To claim otherwise is to subscribe to the fallacious assertion that our constructs are *not* constructs (that is, terms invented by us to capture what we think we observe: for example, intelligence, memory, anxiety, dogmatism), but are real entities to be found in the phenomenal world (could the skull be appropriately pierced). Direct physical and sensory contact with our constructs is plainly impossible, so it must be owned that the idea of a construct's validity cannot imply a secure anchoring to something independent of human perception and language. Instead, validity implies something about the definitional and explanatory usefulness of the construct in question.

From this it follows that the validity of a construct (or of an instrument) is incomprehensible and valueless until it is couched in a conceptual context. For instance, the validity of an IQ test is meaningless if it is solely a matter of correlating its measures with those of another IQ test, because this strategy regresses to circularity. Meaning comes to the validity of a construct (such as IQ) or a measure (such as an IQ test) when the construct is shown to be connected with other constructs in a fashion which permits useful explanations and testable predictions.

Not surprisingly, this view of the validity of a construct permeates the physical sciences: Planck's *h* and Pauli's neutrino are valid constructs only because they successfully plug conceptual gaps in elaborated theoretical structures. Here it is plain that, for meaning, the concept of validity must incorporate the *place* of a construct within a network of constructs which, together, bear some relationship to observations.

Margenau (1950) puts it this way:

> Constructs admissible in science must be multiply connected; they may not be insular or peninsular; sets forming an island universe must be excluded. (p. 87)

Two significant sorts of connections are, for Margenau, formal and epistemic. Formal connections are those which incorporate constructs within a network of related constructs. Epistemic connections are those which connect constructs to observations. Significantly, constructs not so connected are inadmissible to the sciences because they are impotent for explanatory and predictive purposes.

These considerations can be applied to the multitrait-multimethod validation matrix of Campbell and Fiske. This model incorporates two important features. First, it tests for the convergence of a novel construct or trait with a similar and known construct, and the divergence of the novel construct from a dissimilar but known construct. In the epistemological terms developed above, convergence tests for coherence of the novel construct within a conceptual network. (A convergence represented by a perfect correlation is not expected, of course, for that would imply that the construct is *identical to* rather than *convergent with* another construct.) Divergence assures us that the construct under investigation does not converge with everything and is, therefore, a unique construct. In measurement terms, obtained divergence permits the assumption that obtained variance is not simply the consequence of applying a single measuring technique.

The second important feature of the Campbell and Fiske matrix is the incorporation of more than one method of measurement. In measurement terms the tapping of two *varieties* of behaviour allows one to check the variance that may be a function of the tapping of behaviour in a particular way. The argument from an epistemological viewpoint is equally compelling; the greater the number of different types of epistemic connections demonstrated for a construct, the more sound the construct is taken to be. This is a particularly important feature of test and construct validation in our view, especially when the subject of a validation inquiry is a classroom observation instrument. To deny its significance is to assume that a teacher's classroom behaviour is independent of any mental states (beliefs, knowledge, myths and so forth) which could be "tapped" by other measures.

These epistemological and measurement viewpoints lead to the minimal Campbell and Fiske matrix given in Figure 1.

So long as it is understood that "construct" and "trait" are interchangeable, the conceptual significance of the minimal matrix may be described as follows. Suppose that the validity of a test reputed to measure the novel construct (Test A) by an observation instrument (Method 1) is in question. First, that measure should correlate significantly with Test B, since it measures the same trait, using a different method, a written test (Method 2). This multimethod axis of the model assures us that any validity score of Test

	Novel Construct	Known and Related Construct	Known and Unrelated Construct
Method 1	Test A	Test C	Test E
Method 2	Test B	Test D	Test F

Figure 1. The minimal components of a Campbell and Fiske matrix.

A is not an artifact of the *method* of measurement used. Second, so that we can be assured that A and B measure a construct which is not isolated but coheres with another construct having a defined conceptual coherence, a second construct (known and related) representing a coherent and converging construct is identified, and instruments C and D are selected with the intent of having A and B scores correlate significantly with those of C and D. Third, in order that A and B are measures of a distinctive construct (that is, they don't converge with everything) it is necessary to show that they discriminate from other measures, E and F, of a known construct conceptually understood to have no relationship to the other two constructs. The theoretically derived hypotheses are that measures A, B, C, and D will correlate positively and significantly, while measures E and F, while correlating positively and significantly with each other, will have insignificant correlations with each of A, B, C, and D scores.

Correlational techniques are naturally appealing for analyzing the data in this matrix. Yet, as Kalleberg and Kluegel (1975) have pointed out, the predicted zero order correlations provide ambiguous information. Basically, these correlations do not include partial correlations (interactions) that may exist among the methods and traits measured. The suggestion is that a path analytic approach to the analysis of the multitrait-multimethod matrix be undertaken on the data collected after a confirmatory factor analysis has been performed. While this is not the only approach that could account for interactions, it is a sensible one. One of the requirements of the approach would seem to be a large number of method x trait cells within the matrix in order to reduce error. Many empirical studies, including the one that will be reported in this paper, fail to meet this requirement.

Critique of a Recent Validation Study

In a recent paper, Borich and Malitz (1975) propose to use the Campbell and Fiske model to validate three classroom observation schemes, IAST, FAIR and COS, reported by Rosenshine and Furst (1973). The use of the multitrait-multimethod matrix in this study presents a bit of a puzzle for the following reasons. First, *all* the instruments used are classroom observation instruments, and so the study does not appear capable of tapping behaviours through more than one method, except in the sense that the same behaviour (a videotape of teaching) is examined using three observation instruments. Accordingly, the study uses a monomethod technique. Consequently, correlations among measures reported include to some unknown degree, a "single method" component independent of the construct being measured.

Second, the Campbell and Fiske model, as interpreted in the previous section, requires the use of a measure of a construct that is not anticipated to correlate with any of the other constructs or measures — it tests the divergent validity of the construct in question. This aspect of the model is not straightforwardly present in the study for the three observation instruments have very similar features. For example, all three systems have a category equivalent to "lecture." Also, all three have a category for coding the equivalent of "affective moves." Obviously, a coder *cannot* check both "lecture" and "affective move" for the same piece of teacher verbiage on a single scheme. So, use of one category precludes use of another — the data are ipsative. It is erroneous to correlate such data and construe the results as a test of validity. Rather, we have here a strenuous test of reliability. This is plainly revealed in Figure 2, adapted from the study of Borich and Malitz.

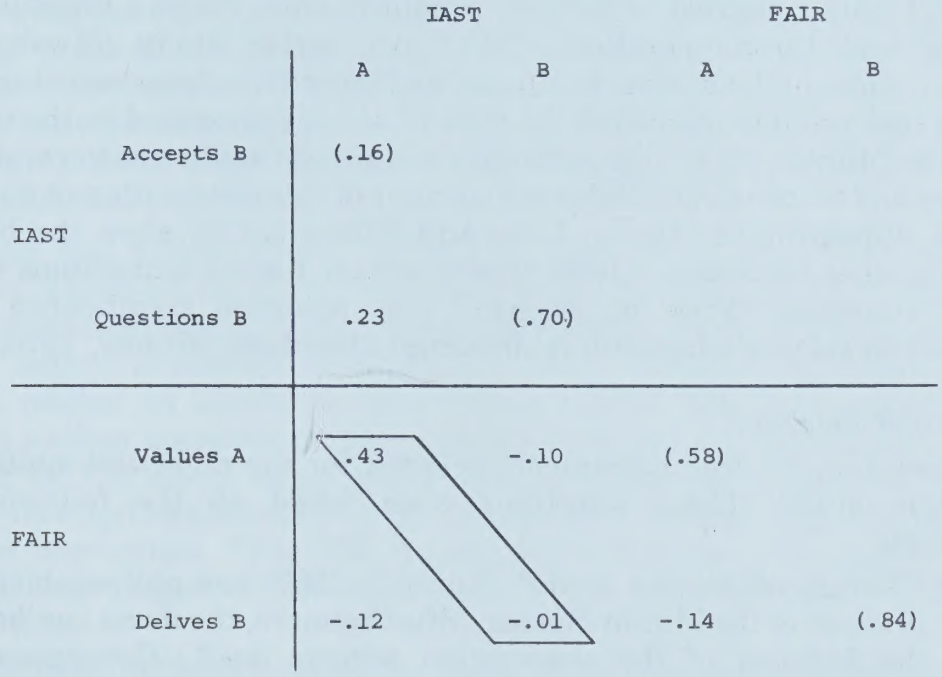

Figure 2. Simplified illustration of the validation model used by Borich and Malitz.

Typically, validation is displayed in the "validity diagonal," (A-FAIR and A-IAST, B-FAIR and B-IAST). Yet it is to be remembered that A and B scores are ipsative so correlations within the validity diagonal are reflections of coding error and not of convergent or discriminant validity.

This misinterpretation of the model can explain Borich and Malitz's nonplus that:

> When one pairs categories, one is hypothesizing that the two categories measure the same behavior, that is, they will demonstrate convergent validity. But at the same time, one is hypothesizing that each of the paired categories differs from other categories in its own system and in the other system. In other words, a hypothesis about convergent validity necessarily includes a hypothesis about discriminant validity. (p. 430)

Of course! And it is precisely this logical entailment which signals that the Campbell and Fiske model is inappropriately used here.

The last problem is implicit in the above quotation in which the talk is of categories *measuring* behaviour. In fact, categories record *counts* of

behaviour, so when a category is used it is not measuring a construct or a portion of one, except when a construct is mounted atop the frequencies, conceptually; instead it is *recording* a behaviour. Since in the Borich and Malitz study, constructs and devices purporting to tap them are not in evidence, the notion of construct validity simply does not apply, much less that of convergent and discriminant validity.

Convergent and Discriminant Validation of "View of Science"

It has been the overall argument of this paper thus far that the Campbell and Fiske model is especially suited to validating classroom observation instruments whose measures allegedly have epistemic connections to constructs. It is the purpose of the remainder of this study to instance that this is so and, by this, to establish the validity of the construct "View of Science." This construct is derived primarily from Nagel's descriptions of Realism and Instrumentalism (1961). An earlier study (Munby, 1973) analyzed these philosophical positions, and reported a classroom observation scheme that reliably identified the view of science presented in the teaching discourse (Munby, 1974). The purpose of the present study was to validate this measure and its construct. Relevant portions of this observation or analytical scheme, appearing in Munby, Kitto and Wilson (1976), show the basis for selecting other constructs which appear to have formal connections with the bipolar construct "View of Science." The potential significance of this construct to science education is discussed elsewhere (Munby, 1976).

Instrument Selection

Figure 3 shows the instruments selected for the multitrait-multimethod validation study. These selections were based on the following considerations.

The "Nature of Science Scale" (Kimball, 1967) has philosophical bases similar to those of the Munby System. (Furthermore, the items can be seen to reflect the features of the observation scheme itself.) Conceptually, the "Nature of Science Scale" and the Munby System measure the same trait, using different methods. That is, it is assumed that a teacher's view of science and his teaching of science are similar, and that "View of Science" is epistemically connected both to teaching behaviour and to responses to the "Nature of Science Scale."

		CONSTRUCTS (TRAITS)		
		View of Science	Authoritarianism (convergent)	Performance Anxiety (divergent)
METHODS	Observation	Munby System	Flanders System	Behavioral Checklist
	Written	Nature of Science Scale	California-F Scale	PRCS

Figure 3. Instruments used in the validation study.

Analysis of the social implications of Realism and Instrumentalism (Munby, 1973) suggests that one whose posture is Realist would evidence authoritarianism and dogmatism. For proponents of Realism, science provides us with unalterably true statments about the world, and science is the only way of obtaining accurate and reliable information about the world — this being recognizably similar to Roszak's "Myth of Objective Consciousness" (Roszak, 1969) which, for him, underlies society's posture toward science. These and statements characterizing Realism appear to reflect an authoritarian and dogmatic view about truth and about science. Instrumentalism, on the other hand, makes softer claims about truth, the nature of theories, and the capability of science. So, an Instrumentalist is expected to be neither authoritarian nor dogmatic.

This conceptual cohesion is cemented by Rokeach's work (1960), in which the following are associated with authoritarianism and dogmatism: concrete thinking, intolerance of ambiguity, premature closure of perceptual processes. Such characteristics are consistent with a Realist view of science, and inconsistent with an Instrumentalist view. Accordingly a formal (conceptual) connection is posited between View of Science and Dogmatism, though they are not identical constructs.

Dunkin and Biddle (1974, pp. 93-101) trace the origin of the Flanders system of classroom interaction to earlier work in which teaching behaviour was characterized along the continuum democratic — autocratic. Additionally, they report a number of studies in which a teacher's "directedness" appears related to scores on dogmatism scales. For these reasons, the Flanders system appeared a likely candidate for the convergent observation measure. Of the many ratios which can be calculated using the Flanders system, the Pupil Initiation Ratio (PIR) was judged closest conceptually to the construct dogmatism, "The PIR is proposed to indicate what proportion of pupil talk was judged by the observer to be an act of initiation" (Flanders, 1970, p. 102).

It was found difficult to identify a distinctively divergent trait having both a written measure and an observation measure suitable for classroom observation for, by definition, the construct "View of Science" represents a basic philosophical posture toward science and reality and so it can be expected to converge somewhat with several attitude and knowledge measures. (A search of the literature revealed that there are very few constructs for which more than one method of measure is available.) A potentially suitable candidate was found in the work of Paul (1966). His work on performance anxiety includes two measures: a written Personal Report of Confidence as a Speaker (PRCS), and a timed observation checklist for performance anxiety in speaking. Theoretically, there is no reason to suspect formal connections between performance anxiety and "View of Science" or Dogmatism.

Given the amount of work involved in coding classroom behaviour and the availability of at least three measures of dogmatism, a pilot study was designed to test the written method portion of the Campbell and Fiske model (Munby, Kitto & Wilson, 1976). An additional "trait" of random numbers was added to the validation model. The Nature of Science Scale correlated ($r = .51$, $p < .001$) with the California-F Scale, and the PRCS and Random Numbers

had near zero correlations with the Scales, as predicted. Accordingly the California-F Scale was selected for the validation study.

Hypotheses of the Validation Study

The hypotheses of this multitrait-multimethod study can be represented on a validation matrix, as in Figure 4. The validity diagonal is represented by AB. Anticipated significant positive correlations in the matrix are represented by "+" while near-zero insignificant correlations are represented by "NS." The correlations along the diagonal CD will be measures of reliability. A total of 21 correlations is predicted.

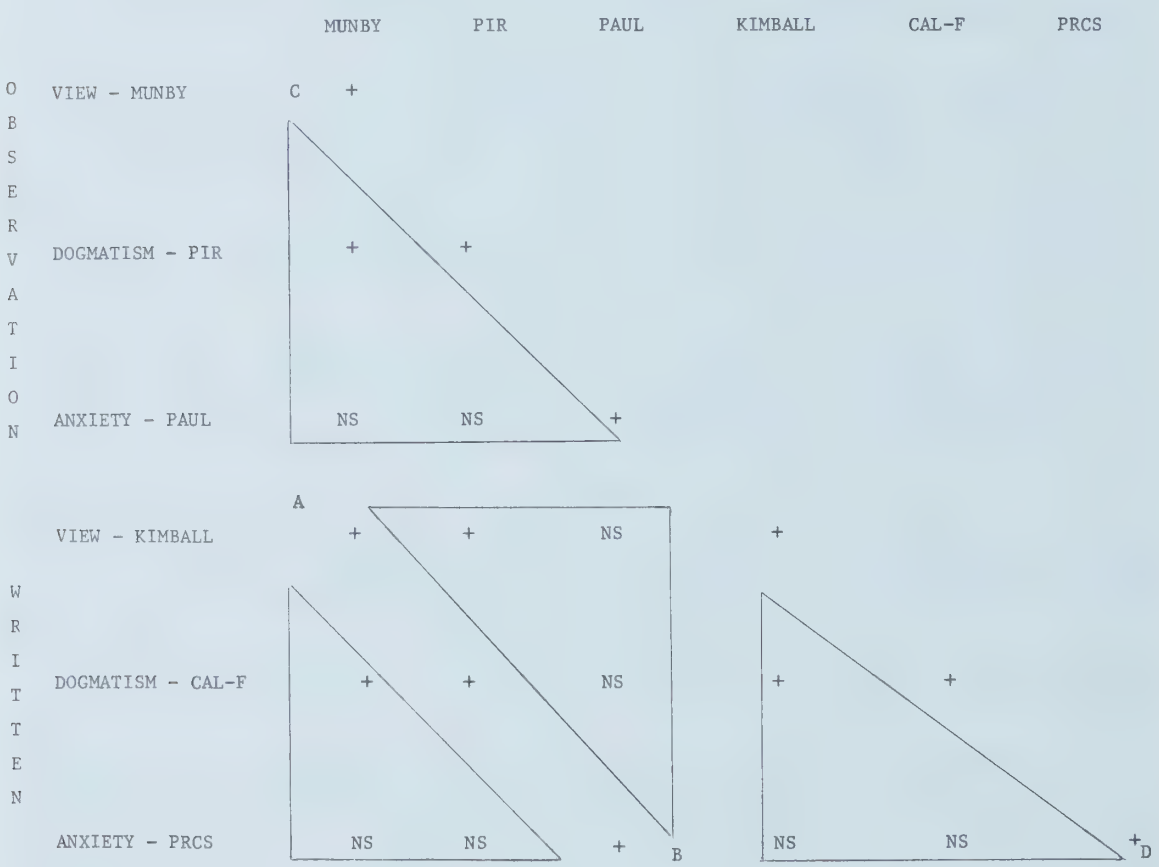

Figure 4. Hypotheses of the validation study.

Methodological Considerations

The scaling of the "Nature of Science Scale" and the PRCS were changed from True/False to Likert-type (+3, +2, +1, -1, -2, -3) to correspond to the California-F Scale.

Items from the three written measures were randomized and printed as a single composite test to ease administration and to attempt to reduce effects of response sets.

The Munby System was used to judge each 30 seconds of teaching discourse as "Very Instrumental" (2), "Moderately Instrumental" (1), "Intermediate" or "Marginal" (0), "Moderately Realist" (1), "Very Realist" (2). A "Realist Ratio" (the total R score divided by the total R and I scores, represented as a percentage) can then be computed for each lesson.

A computer program was devised to produce Flanders matrices, and to compute values of PIR (Kitto, 1976).

Subjects

Fifty-five science teachers randomly selected from local schools, agreed to participate in the study. Each was asked to provide the time and date of a class which would involve teaching, as opposed to testing or conducting laboratory work. While participants were informed that the research was to validate an instrument, the nature of the variables and instruments was not disclosed until the data were collected.

Each lesson was recorded on audiotape for later analysis by the Flanders and Munby Systems. Live observation was used to produce a PRCS score, and in 25 of these cases, two observers were present to provide reliability data. After a lesson was recorded, the teacher completed the composite test.

The lessons represented a wide range of subjects and grade levels. Eight lessons were removed from the study because they were either too short (less than 10 minutes) or they produced no score on the Munby System. (Despite the requests, some teachers had scheduled laboratory exercises which commenced with managerial matters about performing the exercise.) These exclusions left 47 lessons for analysis.

Data Analysis and Discussion

The program used for producing the multitrait-multimethod matrix was adapted from the SPSS package (Nie, Bent & Hull, 1970). Spearman corrections were made for reliability coefficients and the validity coefficients were corrected for attenuation, yielding the matrix in Table 1. (Reliability values for observation measures should be understood as interobserver agreements.)

Of the 21 correlations in the matrix only three failed to correspond with those predicted in Figure 3: PIR and the Munby System did not correlate significantly; the written measure of anxiety as a public speaker (PRCS) correlated significantly with the California-F Scale, whereas in the pilot study (Munby, Kitto & Wilson, 1976) the correlation was .09 ($p < .13$) and, surprisingly, the written and observation measures of anxiety failed to correlate, suggesting that while each measure functions to show the discriminant validity of the construct "View of Science", anxiety as a public speaker did not meet minimal convergent validity criteria.

The remaining significant correlations, while in the anticipated direction, are not high. Their failure to account for more substantial portions of the variance is not too problematical, though, for two reasons. First, it needs to be recalled that convergence, in this model, will not approach unity simply because this would demonstrate that the converging constructs are the same. In other words, the formal connections tested by the multimethod-multitrait model are just that — *connections*. Second, while the assumption that teaching behaviour is reflexive of teacher views or beliefs is conceptually attractive, it is not very likely that the assumption is thoroughly tested by coding one sample of each teacher's teaching. The revelation that the correlations are significant is cause to view the assumption with respect, and to put it to a fuller and more exacting test.

TABLE 1
MULTITRAIT-MULTIMETHOD MATRIX (N = 47)[a]

		MUNBY	PIR	PAUL	KIMBALL	CAL-F	PRCS
OBSERVATION	VIEW - MUNBY	.97 (.001)					
	DOGMATISM - PIR	.17 (.13)	.99 (.001)				
	ANXIETY - PAUL	-.16 (.14)	-.06 (.34)	.96 (.001)			
	VIEW - KIMBALL	.42 (.002)	.38 (.004)	-.12 (.21)	.69 (.001)		
WRITTEN	DOGMATISM - CAL-F	.30 (.02)	.35 (.008)	-.19 (.10)	.68 (.001)	.73 (.001)	
	ANXIETY - PRCS	-.12 (.21)	-.04 (.40)	-.12 (.21)	-.04 (.40)	.30 (.02)	.89 (.001)

[a] Significance levels in parentheses

Two further matrices were produced to investigate the validity of the construct "View of Science" at the ends and at the middle areas of the distribution of scores on the Munby System, for the middle range was quite flat. A high group of fifteen teachers was identified whose "View of Science" score exceeded 69 ($z > +.43$) and a low group whose score fell below 57 ($z < -.43$). The matrix of this group, with coefficients corrected as before, appears in Table 2. While all values move closer to the anticipated values suggested by Figure 3, the anomalies noted above persist.

Conclusion

This study has attempted to provide an epistemological interpretation of the Campbell and Fiske multitrait-multimethod validation study, and to show its theoretical potential for the validation of classroom observation schemes. This view of the model was used to critique a similar validation study. A validation study on the Munby System and the construct "View of Science" was reported as an instance of a more appropriate use of this validation technique.

Ray and Heeler (1975) opine that a single validation study is insufficient and that a series ought to be undertaken. The results of the present study support that position. Although they are not entirely satisfactory, the results

TABLE 2
MULTITRAIT-MULTIMETHOD MATRIX, HIGH-LOW GROUP (N = 30)

		.MUNBY	PIR	PAUL	KIMBALL	CAL-F	PRCS
OBSERVATION	VIEW - MUNBY	**.99** (.001)					
	DOGMATISM - PIR	.22 (.12)	.99 (.001)				
	ANXIETY - PAUL	-.15 (.22)	0 (.49)	.99 (.001)			
WRITTEN	VIEW - KIMBALL	.50 (.002)	.44 (.01)	-.08 (.34)	.76 (.001)		
	DOGMATISM - CAL-F	.39 (.02)	.49 (.003)	-.19 (.16)	.65 (.001)	.71 (.001)	
	ANXIETY - PRCS	-.11 (.28)	0 (.49)	-.19 (.16)	-.08 (.34)	.32 (.04)	.93 (.001)

show that "View of Science" has promise as a verifiable construct. A further validation study might investigate the appropriateness of the sampling of teaching behaviour and so investigate more thoroughly the assumed relationship between teacher behaviour and teacher belief and view, possibly using the techniques suggested by Kalleberg and Kluegel (1975).

It has been suggested above that few classroom observation instruments are subjected to systematic validation, and it may be that some are not amenable to the procedures reported here. Also notable is the impressive lack of unanimous results for studies using such devices. While it remains unwise to posit a causal connection between these observations, it does seem plain that if research using complex and detailed classroom observation schemes is considered worth the effort, it is likely useful to validate these devices in a less perfunctory fashion.

References

Borich, G. D., & Malitz, D. Convergent and discriminant validation of three classroom observation systems: A proposed model. *Journal of Educational Psychology,* 1975, *67,* 426-431.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 1959, *56,* 81-105.

Dunkin, M. J., & Biddle, B. J. *The study of teaching.* New York: Holt, Rinehart and Winston, 1974.

Flanders, N. A. *Analyzing teaching behavior.* Reading, Massachusetts: Addison-Wesley, 1970.

Kalleberg, A. L., & Kluegel, J. R. Analysis of the multitrait-multimethod matrix: Some limitations and an alternative. *Journal of Applied Psychology,* 1975, *60,* 1-9.

Kimball, M. E. Understanding the nature of science: A comparison of scientists and science teachers. *Journal of Research in Science Teaching,* 1967, *5,* 110-120.

Kitto, R. J. Production of Flanders matrices from 3-second coding of classroom teaching. Unpublished computer program. Faculty of Education, Queen's University, 1976.

Margenau, H. *The nature of physical reality: A philosophy of modern physics.* New York: McGraw-Hill, 1950.

Munby, A. H. The provision made for selected intellectual consequences by science teaching: Derivation and application of analytical scheme. Unpublished doctoral dissertation, University of Toronto, 1973.

Munby, A. H. Munby system. In A. Simon & E. G. Boyer (Eds.), *Mirrors for behavior III: An anthology of observation instruments.* Wyncote, Pennsylvania: Communication Materials Centre, 1974.

Munby, A. H. Some implications of language in science education. *Science Education,* 1976, *60,* 115-124.

Munby, A. H., Kitto, R. J., & Wilson, R. J. Validating constructs in science education research: The construct "View of Science". *Science Education,* 1976, *60,* 313-321.

Nagel, E. *The structure of science.* New York: Harcourt, Brace & World, 1961.

Nie, N. H., Bent, D. H., & Hull, C. H. *Statistical package for the social sciences.* New York: McGraw-Hill, 1970.

Novak, J. D. Science research review series: A summary of research in science education for 1972. *Science education information reports.* Columbus, Ohio: ERIC Science, Mathematics and Environmental Education Clearinghouse, 1973.

Paul, G. L. *Insight vs. desensitization in psychotherapy.* Stanford, CA: Stanford University Press, 1966.

Ray, M. L., & Heeler, R. M. Analysis techniques for exploratory use of the multitrait-multimethod matrix. *Educational and Psychological Measurement,* 1975, *35,* 255-265.

Rokeach, M. *The open and closed mind.* New York: Basic Books, 1960.

Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In R. Travers (Ed.), *Second handbook of research on teaching.* Chicago: Rand McNally, 1973.

Roszak, T. *The making of a counter culture.* Garden City, N.Y.: Anchor Books, 1969.

Trowbridge, L. Research review series — Science paper 9: A summary of research in elementary, secondary and college levels in science education for 1970. *Science education information reports.* Columbus, Ohio: ERIC Science, Mathematics and Environmental Education Clearinghouse, 1972.

The Alberta Journal of Educational Research *Vol. XXIV, No. 2, June, 1978*

YEE-LAY JACK LAM
Brandon University

Anxiety Reduction Correlates of Adult Learners: A Longitudinal Study

This exploratory study attempts to identify: (1) major sources of stimuli in the instructional setting that elicit anxiety of adult learners, (2) the changes of anxiety level through the entire course period, and (3) potent factors that either bring about reduction of anxiety or retain a low level of anxiety. Data from 81 adult learners attending six different professional education courses in the Faculty of Education at Brandon University suggest that there are basically three types of anxiety: anxiety about the course setup, anxiety about the interpersonal process, and anxiety about the evaluative outcomes. Levels of anxiety vary but they all display some degree of decline towards the end of the course. Of the potent instructional factors accounting for reduction in anxiety, there are a few whose effects are persistent throughout the duration of the course, some whose influence on anxiety emerges or re-emerges in different periods of course instruction, and still others whose impact on anxiety are totally displaced towards the end of the course. Changing needs and expectations of adult learners are significantly related to the amount of anxiety experiences. (Dr. Lam is Assistant Professor in the Faculty of Education at Brandon University, currently on leave as Post-doctoral Fellow at the University of California at Los Angeles.)

A question of continuing interest to all adult educators is how adult learners effectively acquire knowledge and skills. While possible answers to this question can be sought in many ways, the exploration of some crucial cognitive disorganizers seems to be a fruitful approach. The negative effect of anxiety, in particular, has attracted the widespread attention of people from various disciplines: neuro-physiologists, biochemists, psychiatrists, and psychologists. Moreover, there is no lack of theories developed to account for the phenomenon regarding anxiety (Fischer, 1970). Surprisingly, however, few of these efforts have been translated into specific suggestions for educators to follow in fashioning the teaching-learning process. Three plausible explanations can be advanced here for speculation.

1. Consensus is still lacking with respect to the conceptual meaning of anxiety. Overconcern for the development of microtheories of anxiety has so

far prohibited the much needed communication and integration (Spielberger, 1966).

2. Parallel to the theoretical fragmentation of theories are the methodological difficulties of measuring emotions which tend to be transitory in nature and are often ignored in psychological research (Spielberger, 1966).

3. Research probing the relationship between anxiety and concept learning has produced inconsistent results, and further expansion of scope in investigation seems urgently needed to clarify the issue.

An examination of recent research in the area of anxiety suggests that the aforementioned weaknesses are partially overcome: authorities of the various disciplines were brought together for dialogue (Spielberger, 1972). The emergence and continued development of Spielberger's anxiety instrument making explicit the distinction between "state" and "trait" anxiety proves extremely useful in assessing, with greater accuracy, such an emotional impact on learning (Meyers & Martin, 1974).

Problem

To adult educators, however, these efforts still leave a wide gap when applying findings to practice. There is a growing acceptance that anxiety is not simply motivational but frequently a cognitive disorganizer, particularly to aged learners (Cattelli, 1961; Schaie, 1968). The main problem initially confronting adult educators is the identification of the origins of stimuli arising from the actual learning setting, rather than from experimental situations (as is frequently reported in psychological research) which trigger state and trait anxiety. Spielberger's distinction between these two types of anxiety, while useful in resolving inconsistency of research results, may be practically insignificant to those involved in adult education. A major reason is that adult learners are not screened on the basis of trait anxiety and the amount of anxiety detected must be the end product of interaction between the two types.

Another question relatively unexplored but crucial to adult educators is: How does the anxiety level of the learners fluctuate throughout the teaching-learning process and what are pertinent factors that bring about reduction of their anxiety level?

In the light of the proposition and some new tentative empirical evidence (Lam, 1976) that adult learners' needs and expectations are not constant but undergo varying degrees of change over a span of time, another related question to be asked is how do needs and expectation changes affect variation in anxiety levels?

Method

Sample

Instructors offering the education extension courses on campus were initially approached. To minimize group differences in anxiety arising either from exposure to different teaching styles or from different duration of courses, two instructors who offered more courses of similar duration were eventually chosen from those who consented to take part. Essentially it entailed six courses in the areas of Special Education and Educational Administration. All of these courses had been scheduled once every week for three hours throughout the same semester.

Data of this study were compiled from 81 out of 85 adult learners who enrolled in these courses. The four excluded from analysis were those who dropped out from the courses at a fairly early stage.

The ages of these learners ranged from 18 to 62 with the mean being 30. Female learners outnumbered male learners in a ratio of almost two to one (53 vs. 28). In terms of occupation, 28 were senior university students taking their final year in the education program, 35 were teachers, 8 were school administrators, and the remaining 10 had professions that were indirectly related to education.

Instruments

In line with the foregoing argument that Spielberger's instrument, while adequately dissecting state-trait anxiety, fell short of capturing contextual stimuli from which observed state anxiety was elicited and was therefore of little pragmatic value to practitioners in adult education, a nine-item questionnaire attempting to assess both the level and contextual sources of anxiety was constructed and used. These nine items were listed as follows:

(1) To what extent are you contented with the general course outline?
(2) To what extent are you contented with the direction that the course is taking up to this stage?
(3) To what extent are you certain that this course meets your needs?
(4) To what extent are you confident of meeting the course's demands?
(5) To what extent do you feel at ease regarding the instructor's general approach?
(6) To what extent do you feel at ease interacting with your instructor?
(7) To what extent do you feel at ease interacting with your classmates?
(8) To what extent do you feel at ease in sharing your past experiences with others in class?
(9) To what extent are you confident of your performance in the course so far?

Each of these items was assessed by five-point scale with point 1 on the scale denoting, for instance, "very discontented," "very uneasy," "very worried," "very uncertain," to point 5 indicating "very contented," "very easy," "very relaxed," "very certain," and the like. Learners were asked to circle a point that best reflected their feelings at that period of time. In view of the positive nature of the items, anxiety scores of learners were indirectly tabulated by consulting the reverse portion of the scale. For instance, if they expressed feelings of ease in interacting with the instructor by circling point 4 in the scale, their anxiety level in this respect was assumed to have a scale value of 1.

Test-retest reliability of this questionnaire yielded coefficients of .70 (between the first and second period), .71 (between the second and the third period), and .67 (between the first and the third period). These seem to be satisfactorily high for the present exploratory study.

A second set of questionnaires measuring content and structural aspects of the teaching-learning process was administered. In total, it consisted of 17 items developed in a previous study (Lam & Wong, 1974). Reliability of the instrument was detected by inter-item coefficients. Values of internal consistency were found to range from .54 to .68. Given the relatively short instrument, the reliability coefficients are not too low. Basically, content

factors assess learners' cognitive ability to follow the course, the extent to which reading materials and assignments enhance their comprehension, and their perceived adequacy of course preparation on the part of the instructor. Structural factors examine: (a) the effectiveness of the formal interpersonal process (teaching strategies) such as lectures or discussions that enhance learners' comprehension of the courses, (b) the extent of harmony in the informal interpersonal process, i.e., approachability of instructor and classmate sociability, and (c) the condition of the intrapersonal process, i.e., the extent of satisfaction with the instructor and with the course that individual learners experience.

These two sets of questionnaires were administered concurrently to the subjects at the beginning (3rd session), midway point (7th session), and end of the course (2nd last session). Such repeated measures attempted to capture changes taking place during the three critical points of time in the teaching-learning process.

Additional biographical data like sex and age which were reported by Cattell and Scheier to be anxiety-related factors (Cattell, 1966) and occupational backgrounds of the subjects which received relatively little attention in research were all included in the present analysis.

Statistical Analysis Procedures

To identify basic types of anxiety emerging from the nine-item questionnaire, one method of factor analysis, namely Principal Factoring with Iteration using Varimax Orthogonal Rotation, was employed (Kim, 1975). The selection of this method was based on such considerations as getting a better estimation of communality, and a simplified theoretical scheme of underlying dimensions, even though in empirical reality the unrelatedness of the underlying factors may not be established. Such an inherent shortcoming could be borne in mind in the interpretation of the analysis.

Stepwise-Multiple Linear Regression Analyses were repeatedly employed on three separate occasions to identify significant independent variables affecting the types of anxiety at different periods of the courses. Multiple t-test comparisons were used to detect differences of anxiety levels of the four types of learners undergoing varying degrees of need changes.

Results

Types of Learning Anxiety

The anxiety scores of each period (beginning, midpoint, and end of the courses) together with the total mean scores were subjected to Principal Factoring Analysis. The pattern of inferred factors emerging from each period seemed to concur strongly with the overall pattern generated by averaging the responses of each learner over the three periods for the factor analysis (Table 1).

It seems that three underlying types of anxiety can be identified. Factor I is accounted for by items assessing stress with respect to course outline, direction of the course, accommodation of one's requirements, and the general approach adopted by the instructor. Factor II is accounted for by items measuring tension regarding interaction with the instructor and classmates, and with experience sharing. Factor III is accounted for by items tapping

TABLE 1
VARIMAX ROTATED FACTOR MATRIX OF ANXIETY VARIABLES — COMMUNALITIES AND EIGENVALUES

Variables \ Period	Factor I				Factor II				Factor III				Overall Communality
	1st	2nd	3rd	Overall	1st	2nd	3rd	Overall	1st	2nd	3rd	Overall	
Discontentment with course outline	.59	.70	.70	.77	-.32	-.35	.24	.18	-.17	-.05	.35	.11	.63
Discontentment with course direction	.75	.71	.62	.84	-.26	-.20	.22	.12	-.17	-.09	.25	.08	.73
Concerns of Course Assignments	.30	.36	.22	.44	-.19	.09	.16	.26	.54	.51	.80	.65	.68
Uncertainty of Meeting Needs	.61	.74	.57	.78	-.23	.34	.22	.21	.06	.01	.36	.17	.68
Worry about Instructor's Approach	.59	.72	.83	.82	.17	-.33	.37	.25	.14	.37	.30	.19	.78
Tension Interacting with Instructor	.17	.17	.28	.19	.54	.56	.83	.90	.19	.14	.21	.02	.85
Tension Interacting with Classmates	.19	.07	.22	.21	.63	.84	.81	.81	.06	-.03	.27	.19	.73
Degree of Reluctance in Experience Sharing	.21	.35	.23	.18	.53	.83	.87	.84	.31	.18	.15	-.04	.74
Lack of Confidence in Performance	.13	.38	.20	.30	.25	.06	.27	.33	.73	.53	.76	.70	.59
Eigenvalue		4.40				1.41				.73			
% of Variance		67.4				21.5				11.1			

learners' concern regarding their ability to cope with course demands and their confidence in performance. In view of the nature of these items, it seems most logical to equate factor I to anxiety about course setup, factor II to anxiety about interpersonal process and factor III to anxiety about evaluative outcomes.

Further examination of these anxiety variables shows a high degree of communality which indicates a high degree of homogeneity of items assessing anxiety. In terms of variance accounted for by the three emergent factors, it seems that anxiety about the course setup explained a greater proportion of overall anxiety felt by the learners than did anxiety about the interpersonal process and evaluative outcomes (Table 1).

A crucial concern of lumping the anxiety data of the six groups of learners for factor analysis is that it may obscure block characteristics. However, in view of the fact that only two instructors were involved, block variations arising from different teaching strategies and related contextual factors were largely controlled. Supplementary ANOVA analysis detecting group variation levels of anxiety for the three identified types at three different points of time under the two instructors indicated that no statistically significant differences existed (Table 2). This points to the relative unimportance of between-group variations as compared with within-group difference in anxiety level.

Further, by plotting the composite scale scores of each type of anxiety over the three points of time, it seems that learners from these six courses on the

TABLE 2
SUMMARY OF ONE-WAY ANOVA OF GROUP DIFFERENCES IN ANXIETY LEVELS UNDER TWO INSTRUCTORS OVER THREE PERIODS OF TIME

Anxiety	Source of Variation	d.f.	Mean Square	*F*
1st Period				
Anxiety I	Between Groups	1	.1587	.86 (N.S.)
	Within Groups	79	.1856	
Anxiety II	Between Groups	1	1.4220	2.98 (N.S.)
	Within Groups	79	.4764	
Anxiety III	Between Groups	1	1.3595	2.98 (N.S.)
	Within Groups	79	.4559	
2nd Period				
Anxiety I	Between Groups	1	.8500	2.36 (N.S.)
	Within Groups	79	.3608	
Anxiety II	Between Groups	1	1.1710	2.90 (N.S.)
	Within Groups	79	.4033	
Anxiety III	Between Groups	1	.2216	.45 (N.S.)
	Within Groups	79	.4893	
3rd Period				
Anxiety I	Between Groups	1	1.8503	3.46 (N.S.)
	Within Groups	79	.5349	
Anxiety II	Between Groups	1	.9973	1.55 (N.S.)
	Within Groups	79	.6423	
Anxiety III	Between Groups	1	.9670	1.45 (N.S.)
	Within Groups	79	.6650	

whole experienced a relatively low level of anxiety, with anxiety about evaluative outcomes persistently higher than anxiety about course setup or anxiety about interpersonal process. In general, however, the three types of anxiety do exhibit some decline towards the end of the courses and such a decline is more obvious in the case of anxiety about the interpersonal process.

Changing Importance of Anxiety-Related Factors

To determine factors in the teaching-learning process as well as personal variables that account for reduced anxiety, nine stepwise-multiple linear regression analyses were executed. The reduced anxiety scores were derived from the composite scores tabulated by summing up the averages of scale scores defined by the previous factor analysis. Thus, the reduced anxiety score for course setup was generated by adding the individual mean scale score of items 1, 2, 4, 5 of the Anxiety Questionnaire, the reduced anxiety score for interpersonal process, from the individual mean scale score of items 6, 7, 8 of the said questionnaire, and the reduced anxiety score for evaluative outcomes from the individual mean scale score of items 3 and 9. The order of entry of course content, structural and personal factors as independent variables, was determined by the size of the variance accounting for reduced anxiety scores. All the 17 variables were included in the nine analyses. For

brevity, only those variables found to attain significance in beta coefficients as well as *F* values were summarized in Tables 3, 4, 5, and 6.

If attention is focused on the appearance and disappearance of variables statistically significant in accounting for the three types of anxiety at three different periods, some very interesting observations can be made. Each type of anxiety, for instance, has its persistent factors (i.e., variables that are important throughout the whole process), emergent or re-emergent factors (variables that appeared in any two periods), and displaced factors (variables that appeared once and were replaced).

(i) *Persistent Factors.* In the case of Anxiety I (anxiety about the course setup) the persistent factors were "understanding" and "expectation change." These imply that increased understanding of the course content and

TABLE 3

STEPWISE MULTIPLE LINEAR REGRESSION ANALYSES ASSESSING THE RELATIVE IMPORTANCE OF CONTENT, STRUCTURAL AND PERSONAL FACTORS TO ANXIETY REDUCTION OF COURSE SETUP IN THREE DIFFERENT PERIODS

Significant Predictors[a]	Anxiety Reduction of Course Setup											
	1st Period				2nd Period				3rd Period			
	B	R^2	df_1/df_2	*F*	B	R^2	df_1/df_2	*F*	B	R^2	df_1/df_2	*F*
Course Understanding	.69	.23	1/80	13.48	.42	.51	4/77	7.33	.39	.56	2/79	8.74
Expectation Changes	-.31	.27	2/79	6.65	-.43	.43	2/79	19.79	-.23	.66	4/77	7.72
Instructor's Approachability	.41	.31	3/78	4.54			-			-		
Course Satisfaction			-		.35	.33	1/80	6.02		-		
Sex			-		.22	.49	3/78	5.09	.21	.63	3/78	4.70
Age									.21	.79	5/76	2.35

[a] All predictors recorded here have beta coefficients and *F* values significant at .05 level or better.

TABLE 4

STEPWISE MULTIPLE LINEAR REGRESSION ANALYSES ASSESSING THE RELATIVE IMPORTANCE OF CONTENT, STRUCTURAL AND PERSONAL FACTORS TO ANXIETY REDUCTION OF INTERPERSONAL PROCESS IN THREE DIFFERENT PERIODS

Significant Predictors[a]	Anxiety Reduction of Interpersonal Process											
	1st Period				2nd Period				3rd Period			
	B	R^2	df_1/df_2	*F*	B	R^2	df_1/df_2	*F*	B	R^2	df_1/df_2	*F*
Instructor's Approachability	.49	.28	2/79	8.46	.41	.14	1/80	3.93	.42	.52	2/79	8.25
Age	.31	.36	3/78	8.48	.23	.26	2/79	3.30	.21	.57	3/78	5.66
Classmate Sociability	.29	.20	1/80	5.21			-		.34	.38	1/80	10.46
Lecturing	.44	.43	4/77	10.57			-		.22	.63	6/76	3.32
Grade	.21	.45	5/76	2.28			-					
Course Satisfaction			-						.22	.61	5/77	2.53

[a] All predictors recorded here have beta coefficients and *F* values significant at .05 level or better.

TABLE 5

STEPWISE MULTIPLE LINEAR REGRESSION ANALYSES ASSESSING THE RELATIVE IMPORTANCE OF CONTENT, STRUCTURAL AND PERSONAL FACTORS TO ANXIETY REDUCTION OF EVALUATIVE OUTCOMES IN THREE DIFFERENT PERIODS

Significant Predictors[a]	Anxiety Reduction of Evaluative Outcome											
	1st Period				2nd Period				3rd Period			
	B	R^2	df_1/df_2	F	B	R^2	df_1/df_2	F	B	R^2	df_1/df_2	F
Grade					.24	.08	1/80	4.58				
Expectation Changes	-.39	.20	3/78	11.08	-.33	.16	2/79	6.49	-.36	.37	3/78	12.69
Reading Materials	.31	.15	2/79	5.13			-		.22	.40	4/77	3.84
Classmate Sociability	.21	.37	6/75	2.41			-		.22	.50	6/75	3.64
Discussion	.22	.38	7/74	2.31			-				-	
Occupation									.23	.52	7/74	6.17
Lecturing	.38	.29	4/77	5.95			-				-	
Satisfaction With Instructor			-		.63	.19	3/78	6.52	.45	.43	5/76	5.47
Understanding			-				-		.48	.24	1/80	9.20
Sex			-				-		.22	.33	2/79	3.82
Age			-				-		.21	.56	8/73	4.39
Course Satisfaction			-				-		.22	.57	9/72	2.08

[a] All predictors recorded here have beta coefficients and F values significant at .05 level or better.

great consistency of learners' expectations with those of the instructors were both significant in lowering the level of learners' anxiety about the general arrangement of the courses. In the case of Anxiety II (anxiety about the interpersonal process) the degree of perceived approachability of the instructors and age of the learners are persistent factors. Evidently, friendliness and acceptance by the instructor reduced the tension of interaction in the classroom setting. And, within the age range of the sample, the older the learners, the more relaxed they felt towards the learning process. Richer working and life experiences that accompany age seniority seem to allow older learners to contribute more in formal discussion and to be more at ease in interacting with others in these professional courses.

In the case of Anxiety III (anxiety about evaluative outcomes), it was found that once again little need and expectation changes on the part of learners had a persistent influence. Most probably, little modification and readjustment of one's initial needs or expectations keep the anxiety level about evaluative outome low, though on the whole adult learners still cherish some degree of apprehension about being evaluated.

(ii) *Emergent/Re-emergent Factors*. With respect to those emergent or re-emergent factors, it was found that sex accounted significantly for anxiety about the setup in the second and last period of the evaluative process. Female learners in general maintained a lower level of anxiety than the male as the learning process continued, though such feeling was not marked for both groups of learners in the beginning.

Classmate sociability, lecturing, and grade achieved through course performance were re-emergent factors for reduced anxiety about the

interpersonal process. Apparently, the immediate acceptance by classmates, the perceived understanding of the course through lecturing and some initial formal feedback about learners' performance in the course which were of great importance in reducing learners' initial anxiety about the interpersonal process become once again vital factors towards the end.

Emergent factors for reduced anxiety about evaluative outcomes include grade, reading materials, classmate sociability, occupation, and satisfaction with the instructors. The importance of formal feedback about learners' performance in the course in relation to reduced anxiety about evaluative outcomes in the second period of the learning process, and its disappearance in the last period of the learning process seem to suggest that once adults are assured of their ability to perform well, their self-confidence is reestablished. It was also found that this did not require continued encouragement from the instructor in the last critical stage of learning. On the other hand, the perceived understanding of the course content through assigned reading materials, and the perceived acceptance by fellow learners were vital in reducing anxiety about evaluative outcome. In line with the previous observation, this seems to suggest that the sources of adult self-confidence in performance had shifted from the formal assessment to their own perceived abilities and rapport with classmates.

Satisfaction with instructor, developed in the second stage of the teaching-learning process, continued to exert its anxiety-reduction effect in the last part. This lends support to the assertion that intrinsic contentment with the instructors as individuals serves as a constant check in minimizing anxiety about course evaluation. That occupation of the learners should become an important factor in anxiety about outcome in the last period of the learning process suggests a variation of priorities or values arising from the learners' career background. To learners assuming teaching or administrative capacities, course-taking is one of the many ways of professional development. The performance in the course is, therefore, not as much a concern to them as to university seniors taking courses for degree and career preparation.

(iii) *Displaced Factors.* Turning to the displaced factors of the three types of anxiety, it was found that instructor approachability, important in reducing anxiety about setup in the beginning period, gave way to course satisfaction and to the increased understanding of the course through discussion midway through the course. It would seem that adult learners' initial need to be accepted by the instructors is no longer their major concern as the course develops.

In the case of reduction of anxiety about the interpersonal process, grade, important in the initial stage, was displaced by such new factors as course satisfaction and comprehension of course content through reading material. This implies that the initial attempts to impress instructors through performance in the course became less a concern to adult learners than a development in the genuine delight with and understanding of the course.

Factors accounting for reduction of anxiety about evaluative outcomes such as understanding through discussion and lecturing, which were prominent in the first phase of learning, were replaced subsequently by course satisfaction, sex, age, and general understanding of the course. This

TABLE 6
SUMMARY OF FACTORS RESPONSIBLE FOR THREE KINDS OF ANXIETY REDUCTION AT DIFFERENT PERIODS OF OBSERVATION (ALL VARIABLES ATTAINING SIGNIFICANT *F* VALUES)

Anxiety I (Setup)			Anxiety II (Process)			Anxiety III (Performance)		
Period 1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Understanding	Understanding	Understanding	Instructor's approachability	Instructor's approachability	Instructor's approachability	Grade	Grade	Understanding
Expectation change	Expectation change	Expectation change	Age	Age	Age	Expectation change	Expectation change	Sex
Instructor's approachability	Sex	Sex	Classmate sociability		Classmate sociability	Reading Materials	Satisfaction with instructor	Expectation change
	Course satisfaction	Age	Lecturing			Classmate sociability		Reading materials
	Discussion	Classmate sociability	Grade		Sex	Discussion	Satisfaction with instructor	Satisfaction with instructor
					Course satisfaction	Occupation		Classmate sociability
						Lecturing		Occupation
								Age
								Course satisfaction

suggests that the adjustment to the style of formal process no longer becomes a problem of concern with the passage of time. Rather, the last phase of worry and tension about outcomes arises mostly from the characteristics of the learners — age, sex, cognitive ability, and inner satisfaction.

(iv) *Factors Responsible for all Types of Anxiety*. In search of common factors that at one stage or another affect the three types of anxiety, it was found that characteristics of the learners, sex, age, and the perceived sociability of fellow learners, were the underlying persistent variables. Reference to the sample of the study suggests that within the age range of 18 to 62, the older the learners, the less tension or stress they felt toward the general process. In addition, the female learners expressed a slightly higher level of satisfaction and less anxiety in all phases of the learning, particularly the general setup in the last stage ($r = .25$, $p < .05$). Perceived acceptance by and sociability of fellow-learners played a significant role in anxiety reduction in all stages and this was particularly evident in the last period (correlation coefficient with the three types of anxiety reduction were .39, .61, and .29; all were significant at .01).

This tentative dissection of anxiety correlates of adult learners reveals that structural and personal factors act jointly with course content in bringing about anxiety. In terms of the number of factors involved, anxiety about evaluative outcomes is far more complicated than either anxiety about the setup or about the interpersonal process.

Need Changes and Anxiety Reduction

In the light of the findings and discussion in the previous section that degrees of "changes in needs and expectations" were related to different levels of anxiety about the setup and about the evaluative outcome, further examination of the relationship between the process of need changes and anxiety reduction in all phases of the teaching-learning processes was made. Through multiple t-test comparisons, the amount of anxiety reduction experienced by the sampled learners was analyzed, with learners categorized into four groups as in a previous study (Lam, 1976), i.e., in terms of their initial similarity/ dissimilarity of needs with instructors and their ultimate professed great/small need changes.

Significant differences in the amount of anxiety about course setup were found between Group II (dissimilar initial needs and great final expectation changes, $N = 32$) and Group I (dissimilar initial needs and small final changes of expectation, $N = 10$) and also between Group II and Group IV (similar initial needs and small expectation changes, $N = 28$) in the initial stage of learning process (t values were 2.14 and 2.52 respectively, $\overline{X}_{II} = 2.45$, $\overline{X}_{I} = 2.10$, $\overline{X}_{III} = 2.15$, $p < .05$ and above). These reflect that those who cherished a set of expectations different from the instructors and subject more readily to changes through the teaching-learning process had a higher initial tension than those who embraced dissimilar expectations and were not too receptive to new stimuli in the teaching-learning process. The former group's anxiety level was also considerably higher than the group having similar expectations with the instructors and experiencing little adjustment in their expectation. On the other hand, the amount of anxiety expressed by Group II was not significantly different from Group IV who shared similar expectations with the instructors and underwent great changes of expectations in the process, though this last group also displayed (in terms of anxiety means) a higher anxiety level than Group I and Group III (similar initial needs and great final expectation changes, $N = 11$).

Midway through the teaching-learning process, the amount of anxiety level regarding the setup was again detected to be significantly higher in Groups II and IV than I and III (t value between Groups II and I is 2.04, $\overline{X}_{II} = 2.35$, $\overline{X}_{I} = 2.06$, $p < .05$; t value between Group IV and III is 2.77, $\overline{X}_{IV} = 1.94$, $\overline{X}_{III} = 1.54$, $p < .01$). Apparently those who experience greater changes of expectation in the process still cherish a higher level of anxiety about the setup than those who undergo smaller changes.

The amount of anxiety level, in other aspects at all three stages, was found to be insignificant among the four groups.

Tentatively, such evidence suggests that changes of expectations elicited some amount of anxiety about the learning setup, and that anxiety level apparent in the first two periods began to level off rapidly towards the end of the courses so that essentially there were no significant differences among the four types of learners identified towards the end of the courses.

Conclusion

The present study has added to our understanding of adult learners' anxiety states as explained in the context of other research (Meyers & Martin, 1974; Spielberger, 1966), and about their stimuli in the teaching-learning process.

Inherent in the teaching-learning situation are three types of anxiety: anxiety about course setup, about interpersonal process, and about evaluative outcomes. While course setup tends to constitute the predominant source of anxiety, adult learners seem most susceptible to anxiety concerning their learning outcomes. Stimuli of these three types of anxiety vary considerably and reside in many aspects.

From the perspective of duration and stages of influence, anxiety reduction factors fall conveniently into "persistent," "emergent or re-emergent," and "displaced" categories. Both personal factors, age, and changing needs and expectations, as well as variables operationalized from the paradigm of Bergevin and his associates (Lam & Wong, 1974), for instance, course comprehension, instructor's approachability, classmate sociability, course satisfaction, etc., turn out to be important factors in explaining adult learners' anxiety. These findings are not inconsistent with those of Boshier (1973) using a "Congruence" model developed from self-theory.

Relationship between need changes and anxiety reduction reveals that those who initially had expectations different from those of the instructors generally had a higher level of anxiety about the setup. This still holds true for those who were undergoing great expectation transformation midway through the teaching-learning process. Variation in the level of all types of anxiety, however, disappeared for all groups when the end of the courses approached.

Of immediate implication to adult learners, it seems, is the recognition that tailoring course materials to the capability of learners and establishing good human relations between instructors and learners, as well as among learners themselves, are two crucial aspects of anxiety reduction. Given that those undergoing greater changes in needs and expectations experience greater anxiety about course setup in the first two stages of the learning process, it is perhaps pertinent for educators to assist learners to develop more realistic educational objectives at the outset so as to reduce unnecessary tension in the subsequent learning process.

References

Boshier, R. Educational participation and dropout: A theoretical model. *Adult Education,* 1973, *23* (4), 255-282.

Cattell, R. B. Anxiety and motivation: Theory and crucial experiments. In C. D. Spielberger, (Ed.), *Anxiety and behavior.* New York: Academic Press, 1966.

Cattell, R. B., & Scheier, I. H. *The meaning and measurement of neuroticism and anxiety.* New York: Ronald Press, 1961.

Fischer, W. F. *Theories of anxiety.* New York: Harper & Row, 1970.

Kim, J. O. Factor analysis. In N. H. Nie et al., *Statistical package for the social sciences* (2nd ed.). New York: McGraw Hill, 1975, 468-514.

Lam, Y. L. J. Transitional analysis of adult learners' needs. *The Alberta Journal of Educational Research,* 1976, *22* (1), 59-70.

Lam, Y. L. J., & Wong, A. Attendance regularity of adult learners: An examination of content and structural factors. *Adult Education,* 1974, *24* (2), 130-142.

Meyers, J., & Martin, R. Relationships of state and trait anxiety to concept-learning performance. *Journal of Educational Psychology,* 1974, *66* (1), 33-39.

Schaie, K. W. *Theory and methods of research on aging.* Morgantown, W. Va.: West Virginia University, 1968.

Spielberger, C. D. (Ed.) *Anxiety and behavior.* New York: Academic Press, 1966.

Spielberger, C. D. (Ed.) *Anxiety: Current trends in theory and research. Vol. I.* New York: Academic Press, 1972.

The Alberta Journal of Educational Research *Vol. XXIV, No. 2, June, 1978*

J. R. KIRBY
University of Newcastle
and
J. P. DAS
The University of Alberta

Skills Underlying Coloured Progressive Matrices

Raven's Coloured Progressive Matrices and a battery of ability tests were administered to a sample of 104 grade four boys. The relationships between two previously identified subscales of the Raven's and the ability tests and factors were investigated. Even though previous research had identified one of the Raven's scales as a Reasoning by Analogy scale, the present results demonstrate moderately strong relationship (r = .42) between this scale and spatial ability. This relationship is similar to that between the other Raven's scale, pattern completion, and spatial ability. Results are interpreted to indicate that, although subjects could employ several strategies in solving the analogy items, those in the present study used primarily a spatial strategy and to a lesser extent, reasoning; in this regard the analogy subscale was not different from the other subscale. The possible implications of these strategies for test score interpretation and future research are discussed. (Dr. Kirby is with the Faculty of Education at the University of Newcastle in Australia; Dr. Das is Director of the Centre for the Study of Mental Retardation at the University of Alberta.)

Raven's Standard Progressive Matrices and Coloured Progressive Matrices have become popular tests of intellectual capacity. The standard version is suitable for adults and adolescents whereas the coloured version is meant for use with children and the elderly. The Progressive Matrices were originally proposed in 1936 by Penrose and Raven as a series of "perceptual tests" in the *British Journal of Medical Psychology* (cited in Cattell, 1940). Since then the Matrices have been regarded as a test of reasoning or Level II ability (Jensen, 1970). It has been considered by some (Cattell, 1971) as a prototype of a culture-fair measure of intelligence, independent of acquired knowledge and verbal skills. Other researchers stress that it is a test which taps a complex set of processes, some or all of which may be influenced by acquired skills (Vernon, 1950; MacArthur, 1975; Das, Kirby & Jarman, 1975).

The Coloured Progressive Matrices (sets A, Ab, B) (Raven, 1956) has recently been factor analyzed by Corman and Budoff (1974a, 1974b) in a variety of populations in order to discover which subscales, and perhaps by inference which subskills, underlie the test. In their research, they have demonstrated the existence of four subscales or problem types; Discrete Pattern Completion, Simple Continuous Pattern Completions, Continuity and Reconstruction of Simple and Complex Structures (CRS), and Reasoning by Analogy (RA). The first three of these scales involve essentially pattern completion, while the fourth consists of a number of nonverbal analogy items. The test was factorially invariant across ability and cultural groups. It is interesting to note that their empirical classification of problem types corresponds closely with one developed by Egorova.[1]

Egorova, working at the Institute of Defectology in Moscow, has analyzed the skills required for solving the Coloured Progressive Matrices based on her observations of children's performance. In her classification, the following skills are recognized: identity in simple and complex drawings, symmetry in axes, analogy, and miscellaneous skills. While hers is not identical with that of Corman and Budoff, all of Egorova's symmetry in axes items are included in the CRS scale, and the analogy (RA) scales are identical if two items classed as miscellaneous are included. There is reasonable agreement between Egorova and Corman and Budoff on the scales underlying the Coloured Matrices.

The purpose of the present study was to examine the relationships of two of these subscales (CRS, RA) with a variety of ability tests and factors. These relationships would in turn indicate the nature of the skills and processes needed for success on these subscales. More specifically, it was anticipated that CRS would be primarily related to tests of spatial ability, and RA to tests of inductive reasoning.

Method

Subjects

The subjects were 104 normal grade 4 males from urban schools in Edmonton, Canada. Ages ranged from 102 to 127 months, with a mean of 110 months and SD of 4.9 months.

Tests and Procedure

The following seven tests were administered to all subjects in class size groups.

1. *Raven's Coloured Progressive Matrices* (Raven, 1965)
2. *Figure Grouping.* This is a test of reasoning adapted from the Science Research Associates (SRA) Primary Mental Abilities Kit (Thurstone & Thurstone, 1962). The subject is required to select which of four figures is unlike the others and indicate his answer on an IBM answer sheet.
3. *Word Grouping.* Another reasoning test adapted from the SRA kit, this test requires the subject to select which of four words is unlike the others and indicate his answer on an IBM answer sheet.
4. *First and Last Names.* This test was selected from the French, Ekstrom, and Price (1963) Kit of Reference Tests for Cognitive Factors. It is a test of associative memory which requires the subject to study a list of 15 first names

paired with last names for three minutes. Subjects then have two minutes to write as many of the first names beside the appropriate last names which appear on a second sheet of paper in scrambled order.

5. *Word-Number.* This test is another test of associative memory from the French et al., (1963) kit. Subjects study a list of 15 numbers paired with words for three minutes and then have two minutes to write as many of the numbers beside the appropriate words which appear on a second sheet of paper in scrambled order.

6. *Spatial Relations.* This test of spatial ability is adapted from the SRA kit and requires the child to select which of four shapes, when added to a given shape, will form a square. Answers are placed on an IBM answer sheet.

7. *Card Rotations.* This test of spatial ability is adapted from the French et al., kit. Subjects are required to indicate whether given figures have been rotated within the same plane or flipped. Answers are placed on an IBM sheet. Because of the difficulty of this task (112 problems in four minutes), guessing is penalized by subtracting the number incorrect from the number correct.

Results

Correct responses for the two subscales of the Raven's were totalled. The actual items chosen for each scale were selected on the basis of the work of Corman and Budoff (1974a, 1974b) and Egorova.[1] The items composing the CRS scale were: Ab4, Ab5, Ab6, Ab7, Ab8, Ab9, Ab10, Ab11, B3, B4, B5, B6, and B7. The item composing the RA scale were Ab12, B8, B9, B10, B11, and B12. The other two scales of Corman and Budoff were not examined due to insufficient variance in this sample. KR-20 reliabilities for the CRS and RA scales were, respectively, .691 and .777. The remaining six tests were submitted to a principal components analysis, and three factors were extracted and rotated to an oblique Promax criterion. This analysis may be seen in Table 1. The factors were identified as Spatial, Memory and Reasoning. Correlations were then computed between the two Raven's subscales, the Raven's total score (out of 36), and the six ability tests and their three factors. These correlations may be seen in Table 2.

Discussion

It can be seen from Table 2 that both Raven's scales are related to both spatial and reasoning ability, and that the pattern of correlations are the same for each subscale. The correlation between the two subscales was .482. Thus, contrary to expectations, CRS and RA were not differentially related to spatial and reasoning abilities. These results would seem to indicate that the processes or skills underlying performance on these subscales are similar. Correlations cannot be conclusive, but the lack of a strong relationship between the RA scale and the reasoning factor suggests that these problems were not predominantly solved by the present sample of children with an inductive reasoning strategy. Rather, it would seem that they relied more upon a process akin to spatial ability. The fact that the two scales emerged separately in previous factor analyses (Corman and Budoff, 1974a, 1974b), and that the KR-20 reliabilities of these scales were high in the present study, suggests that the scales do contain different problem types. The present

TABLE 1
PRINCIPAL COMPONENTS ANALYSIS WITH PROMAX ROTATION OF ABILITIES BATTERY (N = 104)

Variable	Factor			h^2
	Spatial	Memory	Reasoning	
Figure Grouping	360	-047	488	541
Word Grouping	-126	-054	1.000	847
First-Last Names	-028	715	282	704
Word-Number	031	940	-192	815
Spatial Relations	949	018	-115	801
Card Rotations	873	004	-023	742
Variance	1.763	1.388	1.298	4.449
% of Total Variance	29.4	23.1	21.6	74.2
Correlations between factors:				
Spatial	1.000			
Memory	.308	1.000		
Reasoning	.560	.328	1.000	

TABLE 2
CORRELATIONS OF TWO RAVEN'S SUBSCALES WITH ABILITY TESTS AND FACTORS (N = 104)[a]

	Raven's Scale		Total Score
	CRS	RA	
Ability Test			
Figure Grouping	.305	.331	.355
Word Grouping	.206	.278	.297
First and Last Names	.277	.238	.285
Word-Number	.125	.011	.073
Spatial Relations	.430	.401	.508
Card Rotations	.321	.315	.402
Ability Factor			
Spatial	.437	.423	.525
Memory	.220	.122	.187
Reasoning	.297	.363	.386

[a] r (.01, two-tailed) = .254

All differences between correlations (CRS-RA) nonsignificant.

results show that both problem types are approached with similar strategies, relying upon similar abilities.

The possibility exists that the RA scale requires a form of reasoning not represented by the inductive reasoning factor. While undoubtedly different to some extent, both require the discovery of some sort of rule. If the RA scale is more reliant upon reasoning than CRS is, the correlations between these scales and the reasoning factor should have been more discrepant. The lack of

discrepant correlations involving the CRS and RA scales and the reasoning factor makes it difficult to identify the RA scale as, exclusively, a reasoning scale.

Reasoning is involved in the solution of the two problem types. For both CRS and RA scales, correlations with the reasoning tests and factor are in the neighbourhood of .3. This shows that even though the inducing of a rule is not the most important determinant of a correct solution, it is involved, at least with some children. In particular, the (nonsignificant) difference in correlations with CRS and RA and the verbal reasoning test of word grouping suggests that a verbal reasoning strategy might be approaching importance. A view of the results from this perspective would suggest that performance on these two scales of the Raven's is not independent of verbal acquired skills. One may speculate that most subjects employ the spatial nonverbal strategy, while some, perhaps those that are most successful, employ an increasingly verbal strategy. This latter point has been made by several researchers (e.g., Bridgeman & Buttram, 1975) to explain the poor performance of some groups on so-called nonverbal reasoning tests: verbal strategies can be very successful in these tasks, but are not always used spontaneously.

The importance of the present study does not involve the determination, in any absolute sense, of the nature of the two Raven's subscales. They seem to reflect primarily two different problem types. Rather, the importance is in the description of these two scales in terms of the processes or skills which underlie them. Raven's (1965) claim that performance on the test, and in particular on the two subscales, is independent of learned skills, especially verbal ones, is true for the present sample of nine-year-old boys, who performed that test without any special training or instructions. It is anticipated, however, that cognitive development would result in increased use of verbal strategies, as would training or instructions (Bridgeman & Buttram, 1975; Guinagh, 1971). Differences in cultural background (attributable to both genetic and environmental causes) are also likely to produce different approaches to the test problems (see general discussion of cross-cultural differences in strategies in Das, Kirby & Jarman, 1975).

These results also offer support for a position taken by Hunt (1975). In devising computer algorithms to solve more advanced Raven's Matrices problems, he discovered two possible strategies: a Gestalt algorithm (similar in nature to the passive, spatially-reliant one described above) and an analytic algorithm (similar to our conscious, rule-inducing one). Hunt concluded that the Matrices test could not be viewed as a measure of intelligence, but rather as an indication of cognitive style. It should be noted that he also makes the fascinating speculation (p. 155), that the first strategy or algorithm is evidence of, in Piagetian terms, concrete operations, and the second of formal operations.

The present results should not be interpreted as either favouring or opposing the use of Raven's Matrices. The test has been shown to be useful, and will likely continue to be so. These results do, however, advise caution in referring to the test as a test of "reasoning"; the ability tapped would seem in general to be a combination of spatial and reasoning abilities. Furthermore, it would seem that the interpretation given to a test score could well depend upon the strategy or processes which the subject uses to solve the problems.

These strategies are learnable; as such, they encourage viewing test results as indicative to a certain extent of what needs to be learned. Clearly, more work is required to specify the nature of these strategies, how they can be taught, and how they are related to general intelligence.

Note

1. Egorova. Personal communication to J. P. Das, November, 1975.

This study was partly supported by Canada Council and Killam grants to the first author and Canada Council Grant S-730056 to the second author.

References

Bridgeman, B., & Buttram, J. Race differences on nonverbal analogy test performance as a function of verbal strategy training. *Journal of Educational Psychology,* 1975, *67,* 586-590.

Cattell, R. B. A culture-free intelligence test I. *Journal of Educational Psychology,* 1940, *31,* 161-179.

Cattell, R. B. *Abilities: Their structure, growth, and action.* Boston: Houghton Mifflin Co., 1971.

Corman, L., & Budoff, M. Factor structures of retarded and nonretarded children on Raven's Progressive Matrices. *Educational and Psychological Measurement,* 1974, *34,* 407-412. (a)

Corman, L., & Budoff, M. Factor structures of Spanish-speaking and non-Spanish-speaking children on Raven's Progressive Matrices. *Educational and Psychological Measurement,* 1974, *34,* 977-981. (b)

Das, J. P., Kirby, J., & Jarman, R. F. Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin,* 1975, *82,* 87-103.

French, J., Ekstrom, R., & Price, L. *Manual for kit of reference tests for cognitive factors.* Princeton, N.J.: Educational Testing Service, 1963.

Guinagh, B. J. An experimental study of basic learning ability and intelligence in low socioeconomic status. *Child Development,* 1971, *42,* 27-36.

Hunt, E. B. Quote the Raven? Nevermore. In L. Gregg (Ed.), *Knowledge and cognition.* Potomac, Md.: Lawrence Erlbaum, 1975.

Jensen, A. R. Hierarchical theories of mental ability. In W. B. Dockrell (Ed.), *On intelligence.* Toronto: Methuen, 1970.

MacArthur, R. S. Ecology, culture, and cognitive development: Canadian native youth. Paper presented at a Meeting of the Canadian Ethnic Studies Association, Winnipeg, Canada, October, 1975.

Raven, J. C. *Coloured progressive matrices: Sets A, Ab, B.* London: H. K. Lewis, 1956.

Raven, J. C. *Guide to the Coloured Progressive Matrices.* London: H. K. Lewis, 1965.

Thurstone, L. L., & Thurstone, T. G. *SRA primary mental abilities.* Chicago: SRA, 1962.

Vernon, P. E. *The structure of human abilities.* London: Methuen, 1950.

RONALD F. JARMAN
University of British Columbia

Cross-modal and Intramodal Matching: Relationships to Simultaneous and Successive Syntheses and Levels of Performance Among Three Intelligence Groups

Cross-modal and intramodal matching ability with visual and auditory content was studied in low, normal, and high IQ fourth-grade children. The 80 children in each IQ group were also administered tests of simultaneous and successive syntheses. Results suggested that simultaneous and successive central processes are involved in modality matching performance. It was further suggested that intellectual competence is not uniquely associated with cross-modal as opposed to intramodal matching ability, but instead is reflected in an interaction between strategic behaviour and task demands. (Dr. Jarman is in the Department of Educational Psychology at the University of British Columbia.)

In recent years, there has been an increase in research on the role of sensory systems in children's information processing (Chalfant & Scheffelin, 1969; Freides, 1974). A common paradigm used in the study of sensory systems is modality matching (Ettlinger, 1967; O'Connor & Hermelin, 1971). A stimulus pattern is presented in one modality, followed by a comparison pattern in a second modality, and the subject is required to judge the equivalence of the two patterns. If the two patterns are presented in different modalities the task is designated as cross-modal, and if the same modality is used for both, the task is known as intramodal. Thus, for the auditory and visual modalities as in the present study, two combinations of intramodal matching (A-A and V-V) and two combinations of cross-modal matching (A-V and V-A) are possible in the design of tasks.[1]

The first question of this study is the extent to which sensory systems, as opposed to central processes, account for ability to match information cross-modally and intramodally. This question has received little direct attention despite the large body of research involving the use of tests of matching ability. Freides (1974) notes that the theoretical assumptions of these studies

may be characterized on a continuum, with the extremes designated as "modal-specific" and "nonmodal." The modal-specific point of view places the emphasis on sensory systems in accounting for individual differences in matching ability, and the nonmodal position assumes that central processes account for performance in these tasks.

Current information processing models which do not rely on the modality matching paradigm for their empirical support generally reflect the nonmodal view of sensory systems. An example of these models is simultaneous and successive syntheses as proposed recently by Das, Kirby and Jarman (1975, in press). In brief, Das et al. suggest that information processing may be described in terms of four units in a schematic model. In the input unit, a stimulus may be presented to any one of the receptors, extero-, intero-, and proprioceptors, and within the exteroceptors to any one of the sense modalities. The input can be presented in a parallel (simultaneous) or serial (successive) manner. The stimulus is registered immediately by the second unit, the sensory register, and may be passed on for processing. The third unit, which is responsible for central processing, has three major components: that which processes separate information into simultaneous groups, that which processes discrete information into temporally organized successive series, and the decision-making and planning component which uses the information integrated by the other two components. Finally, the output unit determines and organizes performance in accordance with the requirements of the task.

For purposes of the present study, the important feature of the simultaneous-successive model is that information may be represented either simultaneously or successively in central processing independent of the form of information presentation. For auditory and visual information specifically, the model suggests that the forms of syntheses involved in processing information from these two modalities are not functionally dependent upon the modalities themselves. Thus, auditory information presented successively can be organized in central processes as simultaneous or successive syntheses. In complement, visual information which is presented simultaneously can be processed by simultaneous or successive syntheses.

If various conversions in information are possible in the course of processing as suggested by the simultaneous-successive model, and these conversions are made primarily according to task demands, the dominant dimensions of individual differences in cross-modal and intramodal matching tasks may have their basis in central processes. Accordingly, if it is demonstrated in exploratory factor analyses that cognitive processes of the form suggested by the simultaneous-successive model are related to performance in the various combinations of modality matching tasks, then because the model operationally defines central processes, the nonmodal theoretical point of view will be supported. Conversely, if simultaneous-successive processes are not found among the combinations of modality matching, but rather, cognitive processes appear to be unique to each task dependent upon the modalities involved, this result will be more consistent with the modal-specific theory of sensory systems.

The second question of this study concerns the relationship of cross-modal and intramodal matching ability to general intelligence. A number of strong

conjectures have been made regarding modality matching tasks as indicators of fundamental processes in intellectual competence. These conjectures have referred almost exclusively to cross-modal forms of matching with the implication that it is the comparison of information between modalities that engages basic intellectual processes. Birch and Belmont (1964), for instance, suggest that "this type of intersensory integration in its more complex and general form may be one of the processes that underlies adaptive behaviour and thus IQ" (pp. 859-860). Much stronger conjectures are made by Jensen (1969), who states that the "conceptually most pure" aspects of intelligence are displayed in cross-modal tasks (Jensen, 1969, pp. 10-11).

Despite these claims, there is relatively little empirical information available on cross-modal matching ability and intelligence (Jarman & Das, 1977). Past research has been almost exclusively correlational in design, and indicates that cross-modal matching does not correlate strongly with psychometric intelligence (IQ), although these correlations are generally statistically significant due to large sample sizes, ceiling effects in instruments, and pooling of groups with different means. Further, the assumption of the importance of cross-modal forms of matching is reflected in a paucity of studies which have used intramodal tests for purposes of comparisons of the relative importance of the two types to intellectual development. Thus, the importance of cross-modal relative to intramodal processes in cognitive competence has not been tested directly.

The present study examines the significance of both cross-modal and intramodal matching ability to intelligence by determining the extent to which each of the possible combinations of these tasks discriminates between three intelligence groups. As related to the first question of this study, levels of performance in these tasks among the intelligence groups are interpreted in terms of simultaneous and successive syntheses.

METHOD

Subjects

A complete survey was conducted on the distribution characteristics of the Lorge-Thorndike IQ data available for a large grade four school population. Three IQ groups were formed on the basis of both verbal and performance scores by setting ranges for sampling. The IQ ranges were 65-85 for the low IQ (LIQ) group, 90-110 for the normal IQ (NIQ) group, and 115-135 for the high IQ (HIQ) group. Eighty subjects were randomly selected within each IQ range. All subjects were chosen from predominately middle-class population areas. School cumulative record cards were checked in order to avoid inclusion of children with identified learning disabilities.

Tests

Cross-modal and Intramodal Matching

Four tests of auditory and visual cross-modal and intramodal matching were used as developed by Jarman (1977). Each of the tests employed the same items in order to control for complexity of information, with variation among the tests dependent only on whether the stimuli were presented auditorily or visually. A "same-different" format was used in all four tests (Goodnow, 1971a, cf. 1971b). This item type requires a dichotomous decision

on the part of the subject: the stimulus is presented in the first modality, and then removed or terminated, following which a comparison is presented in the second modality; the subject is asked if the comparison is the "same" or "different" from the stimulus. A complete list of the items in the tests is given in Jarman (1977).

Auditory-auditory matching. This task involves matching a stimulus pattern of 1,000 cycle tones with a comparison pattern of tones. All tones are of 0.15 second duration with variation in patterns created by short pauses of 0.35 seconds and long pauses of 1.35 seconds. A score of one is given for each of the 30 items in the test.

Auditory-visual matching. The items in this test utilize an auditory pattern as a stimulus as in the auditory-auditory test, but this is matched with a comparison visual display of dots which is shown by a slide projector. For a dot of 1 unit in diameter, a short gap is 0.80 units in length, and a long gap is of 7.17 units in length. Each of the 30 items in the test is scored as one mark.

Visual-auditory matching. This test is the converse of auditory-visual matching; the visual display is the stimulus portion of each item and the comparison section is a set of tones.

Visual-visual matching. This test is comprised of sets of visual patterns, where the first pattern of each set is compared to the second pattern, with a score of one given to each correct response for a test total of 30.

Simultaneous and Successive Syntheses

The tests of simultaneous-successive syntheses were drawn from previous research (Das, 1973; Das, Kirby & Jarman, 1975).

Raven's Progressive Matrices. This is a well-known test involving the selection of a missing element from a set of alternatives to complete a spatial matrix. The matrices in the test consist of geometric patterns. The total score on the test is the number of items correct, with a possible maximum of 36.

Figure Copying. The figure copying test was developed by the Gessell Institute (Ilg & Ames, 1964), and consists of 10 geometric figures which are presented consecutively to the subject for reproduction while they are in full view. A maximum score of two is possible for each item, with a test total of 20.

Graham-Kendall's Memory for Designs. This test was developed by Graham and Kendall (1960). A set of 15 geometric figures, each on a separate card, are presented for five seconds to the subject, and following the presentation of each card, the subject is asked to reproduce the figure. The total score on the test is the number of errors made, as described in detail in the manual, with scores for each item varying from zero for a satisfactory reproduction to three for a reversed or rotated reproduction. The maximum score which would result from consistent errors is 45.

Serial Recall. The serial recall task consists of a set of 24 word lists, with four words in each list. Of these lists, 12 contain words which are semantically similar to one another, and the other 12 contain four unrelated words. The two types of word lists are randomly ordered in the test. Each list is presented by the use of a cassette tape recording, following which the subject is asked to recall the list in the order given. The total score for each list is the number of words in the correct position, with a possible total score of four per list, giving a maximum test score of 96.

Visual Short-term Memory. This test was developed by E. Howarth and J. Brown of the University of Alberta. Each of the 20 items consists of a five-section grid which is presented to the subject for five seconds. Following presentation, a two-second filler task of colour naming is used to eliminate rehearsal. The stimuli are single digit numbers and the subject is asked to reproduce the digits on an empty grid following the filler task. Each digit in the correct location is scored as one mark, giving a possible score of five for each item and a possible test total of 100.

Procedure

The set of four cross-modal and intramodal matching tests were administered in a balanced design within each IQ group in order to compensate for possible transfer effects. A simple four by four Latin square design (Cochran & Cox, 1957) was used with 20 subjects in each IQ group randomly assigned to each of four orders of test administration. The tests of simultaneous and successive syntheses were given in random order to each IQ group. Subjects were tested in random order across IQ groups.

RESULTS

Descriptive Information

Table 1 demonstrates that sampling from three ranges of the IQ distribution for grade four resulted in a higher mean chronological age (CA) for the subjects in the LIQ group than the mean CA for the NIQ and HIQ groups. The CA difference between the groups was statistically significant, $F(2,237) = 134.63$, $p < 0.01$. The effect of this difference on all the other variables was checked in subsequent statistical analyses by using CA as a covariate. It was found that after covarying to create a CA-match design (Ellis, 1969; Heal, 1970), the IQ groups still retained the same similarities and differences on all variables.

TABLE 1
MEANS AND STANDARD DEVIATION OF CHRONOLOGICAL AGE AND IQ FOR SAMPLE GROUPS[a]

	LIQ		NIQ		HIQ	
	M	SD	M	SD	M	SD
Chronological Age	119.76	6.53	112.07	3.86	111.88	3.81
Verbal IQ	77.05	5.44	100.89	4.76	124.63	4.95
Performance IQ	76.78	6.47	102.71	6.97	123.23	5.54

[a]Chronological age given in months.

Also given in Table 1 are the means and standard deviations for the verbal and performance IQ scores. The data show that the groups are quite symmetric in their composition in terms of the normal IQ curve.

Exploratory Factor Analyses

In order to identify the processes involved in each of the cross-modal and intramodal matching tests, the complete test correlation matrix for each IQ

group was subjected to principal components analysis. The factors were rotated by both orthogonal and oblique procedures (Gorsuch, 1974). The two types of rotations yielded virtually identical results, and therefore only the orthogonal rotations by Varimax are reported here. The eigenvalue greater than one criterion was used for the selection of the number of factors, which also yielded comparable results to other procedures, including the scree test (Cattell, 1966).

In the factor matrix for each group, the same set of reference factors, namely simultaneous and successive syntheses, was identified. The emphasis of the present study is on the use of these factors to interpret the processes in the cross-modal and intramodal matching tests, and therefore the rationale for labelling the simultaneous and successive factors is not given here. This rationale is supplied in Das (1973) and Das, Kirby and Jarman (1975).

The factor matrix for the LIQ group is presented in Table 2. Factor I of this matrix has a moderate loading from Auditory-visual matching and strong loadings from Visual-auditory matching, Visual-visual matching, and Serial Recall. This factor is unique relative to the usual simultaneous and successive factors, and it also appears as Factor I in the analyses for the NIQ and HIQ groups. The factor appears to represent verbal coding of information. This discussion will return to the nature of the verbal coding factor following examination of the results for the NIQ and HIQ groups.

TABLE 2
PRINCIPAL COMPONENTS WITH VARIMAX ROTATION: LOW IQ GROUP[a]

Test	I	II	III
Auditory-auditory Matching	051	-031	856
Auditory-visual Matching	407	441	217
Visual-auditory Matching	731	-106	257
Visual-visual Matching	792	207	-091
Raven's Progressive Matrices	-034	651	-294
Figure Copying	067	743	134
Memory for Designs	-124	-787	-064
Serial Recall	692	007	057
Visual Short-term Memory	287	204	687
Component Variance	1.912	1.886	1.438
% Component Variance	36.52	36.02	27.46
Eigenvalues	2.751	1.701	1.246

[a] Decimals omitted from factor loadings.

Factor II of the results for the LIQ group has its dominant test loadings from Raven's Progressive Matrices, Figure Copying and Memory for Designs (the Memory for Designs loading is negative because the test is scored for errors). A moderate loading is also notable for Auditory-visual matching. This factor is simultaneous synthesis, and its definition by the first three tests is identical to previous results (Das, Kirby & Jarman, 1975).

Factor III of the LIQ matrix is apparently successive synthesis. This factor has been defined by Serial Recall and Visual Short-term Memory in previous research. Interestingly however, in this analysis for the LIQ group, Serial Recall loads on the coding factor, rather than the successive factor. The sequential nature of Auditory-auditory matching and Visual Short-term Memory defines the successive factor here.

The factor matrix for the NIQ group is given in Table 3. Factor I of this matrix shows a clearer definition of the coding factor, with moderate loadings from Auditory-auditory matching and Auditory-visual matching.

TABLE 3
PRINCIPAL COMPONENTS WITH VARIMAX ROTATION: NORMAL IQ GROUP[a]

Test	I	II	III
Auditory-auditory Matching	408	109	523
Auditory-visual Matching	432	547	004
Visual-auditory Matching	850	076	237
Visual-visual Matching	779	184	306
Raven's Progressive Matrices	319	643	-019
Figure Copying	110	867	194
Memory for Designs	-377	-642	-071
Serial Recall	141	032	751
Visual Short-term Memory	347	-021	693
Component Variance	2.079	1.929	1.511
% Component Variance	37.67	34.96	27.37
Eigenvalues	3.641	1.517	1.004

[a]Decimals omitted from factor loadings.

Factor II is simultaneous synthesis, and is quite comparable in its moderate loading of Auditory-visual matching to the results for the LIQ group. Factor III is successive synthesis, with Serial Recall and Visual Short-term Memory defining the factor as in previous studies (Das, Kirby & Jarman, 1975). Auditory-auditory matching has only a moderate loading on this factor for the NIQ group.

The factor results for the HIQ group are given in Table 4. Factor I is comprised almost solely of the modality matching tests, indicating a verbal coding process common to all modality combinations.

Factor II of the matrix for the HIQ group is a clearly defined simultaneous synthesis factor, and Factor III is successive synthesis.

Collectively, these results demonstrate the expected simultaneous and successive factors in all groups, and also a first factor that has been designated as verbal coding. The verbal coding interpretation of this factor appears to be the most parsimonious, in light of two points of evidence. The first of these is the difference between two distinct strategy types in modality matching as discussed by Kahn and Birch (1968). The first strategy noted by

TABLE 4
PRINCIPAL COMPONENTS WITH VARIMAX ROTATION: HIGH IQ GROUP[a]

Test	I	II	III
Auditory-auditory Matching	647	249	317
Auditory-visual Matching	713	107	201
Visual-auditory Matching	834	-101	-116
Visual-visual Matching	714	-109	054
Raven's Progressive Matrices	216	729	307
Figure Copying	357	676	-108
Memory for Designs	121	-905	-076
Serial Recall	210	019	882
Visual Short-term Memory	158	017	731
Component Variance	2.390	1.904	1.581
% Component Variance	40.68	32.41	26.91
Eigenvalues	2.574	1.671	1.254

[a] Decimals omitted from factor loadings.

Kahn and Birch is the direct recall of tone sequences; in the present study this was suggested by the loading of Auditory-auditory matching on the successive factor for the LIQ group. The second technique is numerical and verbal coding of information. Kahn and Birch noted that in their study roughly half of their subjects reported this technique as a dominant strategy. In the present study, all of the tests for which the stimulus portion was visual were performed by the same cognitive process by all of the groups. It is probable that visual displays encouraged the subjects in all of the groups to use a coding strategy because of the easier clustering and grouping that was possible in this type of presentation.

The second reason for the designation of this factor as verbal coding is perhaps the strongest, and is based on the results for the Auditory-auditory task. It is evident that the numerical coding of an auditory sequence involves a translation to a new form of cognitive representation: e.g., beep, beep, pause, beep, pause, beep . . . becomes 2, 1, 1, etc. (Lehman & Goodnow, 1972). In contrast, recall of an auditory sequence in its original form is not a translation strategy. The reason that the Auditory-auditory task for the low IQ group loaded on the successive factor, then, may be that the subjects in this group attempted to recall the sequence literally. In contrast, the high IQ group performed the Auditory-auditory task by the same strategy as the other tasks: they encoded all information, auditory and visual, to a common code. Further, recall of this code did not become a successive task, and instead coding formed a separate factor, because it was the accuracy of encoding (Biggs, 1969) and ability to match the code with the comparison stimulus that differentiated individuals, not the retrieval of the code. Thus, the verbal coding factor is defined by the organization of information, not the successive recall of the information.

If this interpretation of the verbal coding factor is correct, then the factor

may be a special case of successive synthesis. The two factors, verbal coding and successive synthesis, may be separate in the analyses because the coding factor defines individual differences in terms of verbal-mediational encoding strategies, and the successive factor defines individual differences in terms of rehearsal and memory retrieval. An examination of the modality matching items suggests that this may be the case; once numerically coded, the quantity of information in most items would be well within the capacity limits of the subjects of the present study. Thus, individual differences in modality matching when a verbal coding strategy is used would be more in the encoding phase. In the successive synthesis tasks on the other hand, organized information is presented to the subjects and performance differences are more likely in the efficiency of rehearsal and retrieval. Because encoding and retrieval are likely independent operations (Brown, 1975), two separate factors could be defined, as has been found in the present study.

Mean Differences in Performance

The means and standard deviations for each of the IQ groups on all 10 variables are given in Table 5. For both the cross-modal and intramodal matching tests, and the tests of simultaneous and successive syntheses, an increase in performance is found generally from the LIQ group to the HIQ group. Scores for Memory for Designs are error scores rather than numbers correct, hence its inverse relationship to IQ.

The strength of discrimination of the cross-modal and intramodal matching tests were assessed both among the four tests and relative to the simultaneous-successive tests by a multivariate analysis of variance (MANOVA) followed by a discriminant analysis (Tatsuoka, 1970, 1971). The

TABLE 5
MEANS AND STANDARD DEVIATIONS FOR SAMPLE GROUPS

Test	LIQ		NIQ		HIQ	
	M	SD	M	SD	M	SD
Auditory-auditory Matching	20.07	3.86	23.05	4.14	24.75	3.86
Auditory-visual Matching	19.13	4.49	23.65	4.07	27.01	3.13
Visual-auditory Matching	24.15	4.47	26.67	3.97	28.47	1.88
Visual-visual Matching	25.01	4.06	27.54	2.57	28.35	1.96
Raven's Progressive Matrices	22.74	4.95	25.96	5.13	31.64	3.17
Figure Copying	14.33	2.02	14.56	2.53	16.74	2.51
Memory for Designs	3.77	3.46	3.17	3.23	1.65	2.05
Serial Recall	61.70	19.19	77.86	17.76	91.05	7.60
Visual Short-term Memory	53.46	19.57	64.01	18.61	69.75	17.56

results of the three group one-way fixed effects MANOVA on the 10 variables indicated significant differences between the groups, $F(20,456) = 10.57$, $p < 0.01$.

The vectors of weights for the discriminant functions are presented in Table 6. It was found that the two functions accounted for 57% of the total variance in the discriminant space, with $\chi^2 = 138.64$, $p < 0.01$ for the first discriminant root and $\chi^2 = 18.07$, $p < 0.07$ for the second root. The results in Table 6 for percentage of common variance indicate that Vector 1 represents almost all of the common variance in group differences relative to Vector 2.

TABLE 6
NORMALIZED DISCRIMINANT FUNCTION WEIGHTS[a]

Test	Vector 1	Vector 2
Auditory-auditory Matching	132	117
Auditory-visual Matching	701	-004
Visual-auditory Matching	015	-017
Visual-visual Matching	163	474
Raven's Progressive Matrices	674	-018
Figure Copying	008	-574
Memory for Designs	167	087
Serial Recall	042	-143
Visual Short-term Memory	-017	015
% Variance	91.40	8.60
Eigenvalues	1.094	0.103

[a] Decimals omitted from weights.

The definition of Vector 1 in terms of tests shows that among the simultaneous-successive tests, Raven's Progressive Matrices demonstrates a strong discrimination effect. The strongest discriminator of all of the tests, however, is Auditory-visual matching. The other three modality matching tests are negligible and quite comparable in their discrimination power relative to the Auditory-visual tests.

DISCUSSION

The first question of this study was whether modality-specific sensory systems or nonmodal central processes account for ability to match information cross-modally and intramodally. This was assessed within separate IQ groups in order to determine any possible variations in modal specificity as a function of intelligence. Within the limitations of exploratory factor analysis, the results suggest that the cognitive processes in modality matching may be defined in terms of the central process constructs of simultaneous and successive syntheses. The factor analyses also suggested that successive synthesis may be differentiated into two types. The first is the form defined operationally in previous research (Das, Kirby & Jarman, 1975), which is based predominantly on individual differences in rehearsal and

retrieval processes; this form retained the general designation of successive synthesis in this study. The second form was designated specifically as verbal coding. In terms of the four modality matching tests, it was found that Auditory-auditory matching consisted of successive processes in the LIQ group, a combination of successive and verbal coding processes in the NIQ group, and mainly verbal coding processes in the HIQ group. Auditory-visual matching consisted of a combination of simultaneous and verbal coding processes in the LIQ and NIQ groups, and verbal coding processes only in the HIQ group. Visual-auditory matching and Visual-visual matching consisted of similar verbal coding processes in all three IQ groups.

Thus, the nonmodal view of ability in modality matching has received more support in this study than the modal-specific position. It does not appear possible to identify a pattern of processes in the factor analyses that have been presented that are systematically related to sensory modalities. Indeed, with increasing levels of intelligence, differential factor patterns among the four tests appear to disappear entirely, to the point where in the high IQ group the same cognitive process accounts for the majority of individual differences in all four modality matching tests.

The second question of this study deals with the role of the matching paradigm as an indicator of intelligence. Many past discussions on this topic have referred specifically to cross-modal forms of matching with the inference that it is the conversion of information across modalities that measures cognitive ability. The present study has demonstrated that these conjectures have some degree of validity, but are too simplified. The Auditory-visual matching task was the most powerful discriminator among all of the tests given to the IQ groups. Its cross-modal counterpart however, Visual-auditory matching, demonstrated a much lower discrimination effect. In addition, this low effect for Visual-auditory matching was comparable to that found for the two intramodal tasks, Auditory-auditory matching and Visual-visual matching. Discrimination of the intelligence groups therefore, was not associated with cross-modal as opposed to intramodal processes. Rather, as suggested by the factor analyses, different forms of syntheses are engaged by various combinations of these tasks, and further, these patterns vary among different intelligence groups. It is more likely then that individual differences in strategic behaviour interacted with the different task demands of the modality combinations to produce different discrimination powers among the cross-modal and intramodal tests. Thus, modality matching tests tap considerable variation in strategic behaviour (cf. Jarman, in press) which has likely contributed to the variation in points of view on the theoretical significance of these tasks.

These interpretations of the results of the present study, in turn, suggest two areas of further research. The first of these is concerned with the popular assumption of modal-specific deficits in learning-disabled children (see Freides, 1974). Future studies using both cross-modal and intramodal tasks may reveal that disabilities are based predominantly in central processes, possibly in simultaneous and successive syntheses or in highly similar constructs. Independent research using other paradigms such as dichotic listening also suggests this possibility (Das, Kirby & Jarman, in press).

A second area of possible research is the use of children's sensory

modalities in instructional design. One consequence of the modal-specific point of view has been to incorporate sensory modality variables in remedial and aptitude-treatment experimental programs (Bissell, White & Zivin, 1971; Lilly & Kelleher, 1973; Magoon & Raper, 1977; Silverston & Deichmann, 1975). This incorporation is likely premature, and for children at the age level of those in the present study, it may be more important for instructional purposes to consider the cognitive strategies and central processes characteristically used in different types of tasks. At younger age levels, on the other hand, sensory modalities may prove to be important instructional variables. It may be conjectured that the role of modalities shifts with increasing age, such that a transition takes place from modal-specific processes to central processes as a function, in part, of increasing use of mediational and cognitive strategies. Longitudinal studies spanning the early years of school would contribute much in investigating this possibility.

Notes

1. The nomenclature commonly used to designate the ordering and type of the stimulus pattern and the comparison pattern is the first letter of each modality name. For example, A-V would designate an auditory stimulus followed by a visual comparison.

References

Biggs, J. B. Coding and cognitive behaviour. *British Journal of Psychology,* 1969, *60,* 287-305.

Birch, H. G., & Belmont, I. Auditory-visual integration in normal and retarded readers. *American Journal of Orthopsychiatry,* 1964, *34,* 852-861.

Bissell, J., White, S., & Zivin, G. Sensory modalities in children's learning. In G. S. Lesser (Ed.), *Psychology and educational practice.* London: Scott Foresman & Co., 1971.

Brown, A. L. Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in child development and behavior.* Vol. 10. New York: Academic Press, 1975.

Cattell, R. B. The Scree test for the number of factors. *Multivariate Behavioral Research,* 1966, *1,* 245-276.

Chalfant, J. C., & Scheffelin, M. A. *Central processing dysfunction in children: A review of research.* NINDS, No. 9, Washington, D.C.: U.S. Department of Health, Education & Welfare, 1969.

Cochran, W. G., & Cox, G. M. *Experimental designs.* New York: John Wiley, 1957.

Das, J. P. Structure of cognitive abilities: Evidence for simultaneous and successive processing. *Journal of Educational Psychology,* 1973, *65,* 103-108.

Das, J. P., Kirby, J., & Jarman, R. F. Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin,* 1975, *82,* 87-103.

Das, J. P., Kirby, J., & Jarman, R. F. *Simultaneous and successive cognitive processes.* New York: Academic Press, in press.

Ellis, N. R. A behavioral research strategy in mental retardation: Defense and critique. *American Journal of Mental Deficiency,* 1969, *73,* 557-566.

Ettlinger, G. Analysis of cross-modal effects and their relationship to language. In C. H. Millikan & F. C. Darley (Eds.), *Brain mechanisms underlying speech and language.* New York: Grune & Stratton, 1967.

Freides, D. Human information processing and sensory modality: Cross-modal functions, information complexity, memory and deficit. *Psychological Bulletin,* 1974, *81,* 284-310.

Goodnow, J. J. Matching auditory and visual series: Modality problem or translation problem? *Child Development,* 1971, *42,* 1187-1201. (a)

Goodnow, J. J. The role of modalities in perceptual and cognitive development. In J. P. Hill (Ed.), *Minnesota Symposium on Child Psychology,* V. Minneapolis: University of Minnesota Press, 1971. (b)

Gorsuch, R. L. *Factor analysis.* Toronto: W. B. Saunders, 1974.

Graham, F. K., & Kendall, B. S. Memory-for-Designs Test: Revised general manual. *Perceptual and Motor Skills,* Monograph Supplement 2-7, 1960, *11,* 147-188.

Heal, L. W. Research strategies and research goals in the scientific study of the mentally subnormal. *American Journal of Mental Deficiency,* 1970, *75,* 10-15.

Ilg, F. L., & Ames, L. B. *School readiness.* New York: Harper & Row, 1964.

Jarman, R. F., A method of construction of auditory stimulus patterns for use in cross-modal and intramodal matching tests. *Behaviour Research Methods and Instrumentation,* 1977, *9,* 22-25.

Jarman, R. F. Level I and Level II abilities: Some theoretical reinterpretations. *British Journal of Psychology,* in press.

Jarman, R. F. & Das, J. P. Simultaneous and successive syntheses and intelligence. *Intelligence,* 1977, *1,* 151-169.

Jensen, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review,* 1969, *39,* 1-123.

Kahn, D., & Birch, H. G. Development of auditory-visual integration and reading achievement. *Perceptual and Motor Skills,* 1968, *27,* 459-468.

Lehman, E. B., & Goodnow, J. Memory for rhythmic series: Age changes in accuracy and number coding. *Developmental Psychology,* 1972, *6,* 363.

Lilly, M. S. & Kelleher, J. Modality strengths and aptitude-treatment interaction. *Journal of Special Education,* 1973, *7,* 5-13.

Magoon, R. A., & Raper, C. C. The effect of sensory modalities on learning. *Contemporary Educational Psychology,* 1977, *2,* 55-65.

O'Connor, M., & Hermelin, B. Inter- and Intramodal transfer in children with modality specific and general handicaps. *British Journal of Social and Clinical Psychology,* 1971, *10,* 346-354.

Silverston, R. A. & Deichmann, J. W. Sense modality research and the acquisition of reading skills. *Review of Educational Research,* 1975, *45,* 149-172.

Tatsuoka, M. M. *Discriminant analysis: The study of group differences.* Champaign, Ill: Institute for Personality and Ability Testing, 1970.

Tatsuoka, M. M. *Multivariate analysis: Techniques for educational and psychological research.* Toronto: Wiley, 1971.

The Alberta Journal of Educational Research *Vol. XXIV, No. 2, June, 1978*

W. DOUGLAS SMART
Prince George District Teachers' Association
and
LLOYD O. OLLILA
University of Victoria

The Effect of Sentence-Combining Practice on Written Compositions and Reading Comprehension

This study investigated the effect of sentence-combining practice on the syntactic maturity of written compositions and on reading comprehension. Students from two grade seven classes were randomly assigned to experimental and control groups. The analysis was based on 13 males and 13 females in each group. The experimental group completed 26 sentence-combining lessons during the 16 weeks of study. The sentence-combining problems were based on O'Hare's (1973, 1975) format with a change in combining signals from upper to lower case letters. Syntactic maturity of written compositions was measured by mean words/T-unit from a 45 T-unit sample of writing. Reading comprehension was measured by the comprehension subtest from the Gates MacGinitie Reading Tests — Survey E. *The experimental group made a significant gain ($p < .05$) in syntactic maturity and a significant gain in reading comprehension ($p < .001$). However, these gains were not significant when compared at posttest with the control group. (Mr. Smart is President of the Prince George District Teachers' Association; Dr. Ollila is at the Faculty of Education at the University of Victoria.)*

In the study of elementary education, the relationship between the productive aspect of our written language (composition) and the receptive aspect of written language (reading comprehension) is of particular interest.

A number of studies (Evanechko, Ollila & Armstrong, 1974; Hughes, 1975; and Kuntz, 1975) have found a positive correlation between measures of syntactic maturity of students' writing and measures of reading achievement. For curriculum and instruction purposes, correlations between measures are of little value. What is important is whether instructional procedures designed to improve one aspect of language arts will also have a positive

effect on other aspects of the language arts. Specifically, will sentence-combining practice designed to improve the syntactic maturity of students' written compositions also contribute to improved reading comprehension? The assumption behind this question is that sentence-combining will give the students practice in manipulating the syntactic system of our language. If students become skilled in this manipulation, they should become more effective in dealing with syntactic structures while writing and reading.

Interest in the effect of sentence-combining practice on students' writing has grown considerably since O'Hare's (1973) study with 73 grade seven students. O'Hare based his study on Mellon's (1969) work, replacing the transformational grammar terminology in the sentence-combining problems with actual word cues. O'Hare's experimental students made significant gains on six factors of syntactic maturity including the mean number of words/T-unit. The experimental group was also significantly superior to the control group at posttest on all six factors of maturity. The experimental group's posttest compositions were also judged to be significantly better in overall quality than the control group's posttest compositions. It is unfortunate that the O'Hare study did not include pretest and posttest measures of reading comprehension, as noted by Stotsky (1975) in her review of the curricular implications of sentence-combining.

In a study using sentence-combining with 100 grade seven students over a period of three months, Combs (1975) reported pretest, posttest, and delayed posttest results. The experimental group's posttest compositions showed significantly more syntactic maturity than either their own pretest compositions or the control group's posttest compositions. The delayed posttest showed an erosion of the experimental group's gains but the experimental group was still significantly more mature than the control group.

Cooper (1975) commented that it has been "convincingly demonstrated that sentence-combining will increase both syntactic maturity and overall quality of writing of upper elementary and junior high school children" (p. 72).

The question remains — will sentence-combining practice have an effect on reading comprehension?

Stedman (1971), in a study with grade four students, attempted to determine the effect of a curriculum teaching syntactic embedding on reading comprehension as measured by a cloze test. In his total experimental group, Stedman found a significant gain in cloze scores over the total control group. In a study involving 94 fifth, seventh, and ninth graders, Fisher (1973) used sentence-combining exercises based on 12 transformations followed later in the course by reversing the process to uncover kernel sentences. The students who were exposed to the sentence-combining practice scored significantly higher on the *Stanford Paragraph Meaning Test.* Hughes (1975), in a study at the grade seven level using sentence-combining, used cloze scores as well as the speed and accuracy subtest from *Gates-MacGinitie Reading Test* to measure reading gain. Hughes reported a nonsignificant trend towards improved reading comprehension. In the Combs (1975) study, a number of reading measures were taken. Despite the significantly better posttest writing of the experimental group, there was no effect on the scores of the comprehension subtest of the *Gates-MacGinitie Reading Test — Survey E.*

However, Combs did report a significant advantage for the experimental group over the control group on a specially constructed reading comprehension measure.

Stotksy (1975) suggested the hypothesis that "enhanced syntactic skills through writing activities leads to improved reading comprehension" (p. 61). Based on the evidence above there is no clear answer yet to this hypothesis. The study reported here is designed to shed more light on the question. This study with grade seven students was designed to determine the effect of 16 weeks of sentence-combining practice on the syntactic maturity of students' written compositions and on their reading comprehension. The study was a 2 x 2 factorial design, with the treatment being one factor and sex the other factor. Pretest and posttest measures of both dependent variables were taken.

Sample

All of the 58 students (30 female and 28 male) in two grade seven classes at Vanway Elementary School were randomly assigned to experimental and control groups. Of the original 58 students, three moved away during the 16-week period. Three girls had to be discarded from the data base to make equal sized cells. The analysis is therefore based on 26 students in the experimental group and 26 students in the control group. No attempt was made to control the outside language experiences of the two groups. The students from both groups participated in their regular language arts activities in their own classes. Both regular classroom teachers who worked with the subjects used the study of novels extensively in their language arts program. One of the classroom teachers had the students keep a writing journal in which they wrote on assigned, guided, or self-selected topics in a scheduled period each week. The assumption is made that the randomization procedures balanced these and other outside language experiences.

Procedures

Each group met with the investigator for two 45-minute periods per week for the 16 weeks of the study. Two periods were lost with each group because of holidays and teacher professional days.

The experimental treatment which was called Sentence-Combining was modelled on O'Hare (1973, 1975). During the 30 periods of the intervention, students completed 26 lessons on sentence-combining. The other four periods were interspersed as catch-up, review, and extension lessons. The sentence-combining problems were presented on mimeographed sheets and were completed in a notebook reserved for Sentence-Combining. The O'Hare (1973, 1975) format of presenting combining signals was used in this study, with one change. In Lesson No. 13, a change was made in the word signals from upper case to lower case letters. This eliminated the tendency of many students to leave the signal word capitalized in the final sentence. The following is an illustrative sentence-combining problem.

Mr. Brown told John SOMETHING.
Only he could decide SOMETHING. (that)
He was going to settle down and do sufficient
work to pass. (when to)

The control group, which also met with the investigator for 30 periods

during the 16 weeks, was called Language Arts Growth and Improvement. The control group's activities consisted of three aspects:

(a) Recreational Reading — including the writing of short summaries and orally sharing these with the class.
(b) *Reading for Understanding — Junior* (S.R.A., 1971) — The students started from their own level and completed 40 cards during the intervention period.
(c) Pattern Writing Practice — examining and writing noun and verb phrases and then using them in sentences.

Measurement

Two dependent variables were measured in this study: (a) the syntactic maturity of the students' written compositions and (b) the students' reading comprehension. As shown by the review of the language development studies (O'Hare, 1973), there has been an extensive search for a measure of the developing ability of students to combine more and more kernel sentences. The mean length of T-units has been shown to be a sensitive measure of this developing syntactic maturity (O'Hare, 1973). In order to measure the syntactic maturity of a representative sample of each of the subjects' free writing, the first 15 T-units from three compositions written by each student were analyzed. The T-unit definition used was that of one main clause plus any subordinate clause or nonclausal structure attached or embedded in the main clause. When one word was omitted from a fragment, the word was supplied and the fragment counted as a T-unit. The mean number of words/T-unit per student was calculated by dividing the number of words in the 45 T-unit sample by 45.

The comprehension subtest of *Gates-MacGinitie Reading Test — Survey E* was used as the reading comprehension measure in this study. Three alternate forms of the test are available. Form 2 was used as the pretest measure and form 3 was used as the posttest. The raw scores of the comprehension subtest were used for analysis.

Results

Syntactic Maturity

The mean words/T-unit for the total experimental group changed from 9.77 at pretest to 10.67 at posttest, while the total control group changed from 10.05 to 10.11.

Table 1 summarizes the results of a two-way ANCOVA with repeated measures on the pretest and posttest measures of syntactic maturity for the experimental group. As indicated in the table, there is a significant difference ($p < .05$) between pretest and posttest mean words/T-unit. The experimental group did make a significant increase on the measure of syntactic maturity of their written compositions.

However, Table 2 summarizes the results of two-way ANCOVA on posttest mean words/T-unit of both groups covarying on their pretest scores. As indicated in the table, there was no significant difference between the experimental and control group's mean words/T-unit at posttest. However, a trend ($p = .098$) in favour of the experimental group is evident.

TABLE 1
PRETEST AND POSTTEST MEAN WORDS/T-UNIT:
EXPERIMENTAL GROUP — STATISTICAL SUMMARY

Source	df	MS	F	p
Between *Ss*	25			
Sex	1	0.098	0.021	0.885
Error	24	4.630		
Within *Ss*	26			
Pre-Post	1	10.661	5.928	0.023*
Sex × Pre/Post	1	7.585	4.218	0.051
Error	24	1.799		
Total	51			

* $p < .05$

TABLE 2
TWO-WAY ANCOVA POSTTEST MEAN WORDS/T-UNIT:
STATISTICAL SUMMARY

Source	df	MS	F	p
Groups	1	6.416	2.85	0.098
Sex	1	4.601	2.04	0.160
Groups × Sex	1	4.243	1.88	0.177
Error	47	2.255		

Reading Comprehension

The mean comprehension score for the total experimental group changed from 27.77 at pretest to 33.27 at posttest. The mean comprehension score for the total control group changed from 28.69 to 33.15.

Table 3 summarizes the results of two-way ANCOVA with repeated measures on the pretest and posttest measures of reading comprehension for the experimental group. As indicated in the table, there is a significant

TABLE 3
PRETEST AND POSTTEST COMPREHENSION SCORES:
EXPERIMENTAL GROUP — STATISTICAL SUMMARY

Source	df	MS	F	p
Between *Ss*	25			
Sex	1	108.164	0.737	0.399
Error	24	146.805		
Within *Ss*	26			
Pre-Post	1	393.225	35.989	0.000003*
Sex × Pre/Post	1	0.048	0.004	0.948
Error	24	10.926		
Total	51			

* $p < .001$

TABLE 4
TWO-WAY ANCOVA POSTTEST COMPREHENSION SCORES: STATISTICAL SUMMARY

Source	df	MS	F	p
Groups	1	11.664	0.40	0.531
Sex	1	23.229	0.79	0.377
Groups × Sex	1	16.625	0.57	0.455
Error	47	29.251		

TABLE 5
PRETEST AND POSTTEST COMPREHENSION SCORES: CONTROL GROUP — STATISTICAL SUMMARY

Source	df	MS	F	p
Between Ss	25			
Sex	1	24.921	0.143	0.708
Error	24	174.032		
Within Ss	26			
Pre-Post	1	258.743	14.001	0.001009*
Sex × Pre/Post	1	19.719	1.067	0.312
Error	24	443.543	18.481	
Total	51			

*$p < .01$

difference ($p < .001$) between the experimental group's pretest and posttest comprehension scores. The experimental group did make a significant gain on the measure of reading comprehension.

However, Table 4 summarizes the results of two-way ANCOVA on posttest comprehension scores of both groups covarying on their pretest scores. As indicated in the table, there was no significant difference between the experimental and control groups on the posttest comprehension scores.

The significant growth for the experimental group (Table 3) is not shown to be significant when compared at posttest with the control group. To examine possible reasons for this, the pretest and posttest comprehension scores for the control group were analyzed. The summarized results in Table 5 show that the control group also made a significant ($p < .01$) gain on comprehension scores.

Conclusions and Discussion

The results show that the sentence-combining practice group made a significant increase in both the syntactic maturity of their written compositions and their reading comprehension. However, neither of the increases is significant when compared at posttest with the control group. The present study has failed to achieve the degree of improvement in students' writing that was achieved in the O'Hare (1973) or Combs (1975) studies. It is important to note that the O'Hare study lasted from October to May while the present study ran only from mid-January to mid-May. The effect of sentence-

combining practice over a longer time may lead to significant differences between the groups.

The improvement of the sentence-combining practice group in reading comprehension was significant. The fact that the control group also made significant improvement in reading comprehension was not unexpected. The sentence-combining practice was only one aspect of a total language arts curriculum. Hopefully, many aspects of that curriculum contribute to improved reading comprehension. The improvement by the control group does not diminish the fact that sentence-combining may contribute to improved reading comprehension. The change in the sentence-combining signals from upper case to lower case seems to have facilitated the students' ability to produce well formed final sentences in the sentence-combining problems.

The present study is limited by the size of the sample, the length of the study, and the fact that both groups were instructed by one teacher. Further research on the effects of sentence-combining practice on syntactic maturity of written compositions and on reading comprehension is necessary before sentence-combining is implemented on a large scale. Moreover, research with larger samples of students and other teachers, conducted over a longer period corresponding more closely to O'Hare's (1973) study, should be conducted.

References

Combs, W. E. Some further effects and implications of sentence-combining exercises for the secondary language arts curriculum (Doctoral dissertation, University of Minnesota, 1975). *Dissertation Abstracts International,* 1975, *36,* 1266-A. (University Microfilms No. 75-21,042)

Cooper, C. R. Research roundup: Oral and written composition. *English Journal,* 1975, *64,* (9), 72-74.

Evanechko, P., Ollila, L., & Armstrong, R. An investigation of the relationships between children's performance in written language and their reading ability. *Research in the Teaching of English,* 1974, *8,* 315-326.

Fisher, K. D. An investigation to determine if selected exercises in sentence-combining can improve reading and writing (Doctoral dissertation, Indiana University, 1973). *Dissertation Abstracts International,* 1974, *34,* 4556A. (University Microfilms No. 74-2586)

Hughes, T. O. *Sentence combining: A means of increasing reading comprehension.* Kalamazoo, Michigan: Western Michigan University, 1975. (ERIC Document Reproduction Services, No. ED 112 421)

Kuntz, M. H. The relationship between syntactic attainment and reading ability in the seventh grade (Doctoral dissertation, University of Pittsburgh, 1975). *Dissertation Abstracts International,* 1975, *36,* 2159A. (University Microfilms No. 75-22, 449)

Mellon, J. C. *Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition.* Research Report No. 10, Urbana, Illinois: National Council of Teachers of English, 1969.

O'Hare, F. *Sentence-combining: Improving students' writing without formal grammar instruction.* Research Report No. 15, Urbana, Illinois: National Council of Teachers of English, 1973.

O'Hare, F. *Sentencecraft — An elective course in writing.* Lexington, Mass.: Ginn and Company, 1975.

Reading for understanding — Junior. Chicago: Science Research Associates, 1971.

Stedman, N. A. The effect of a curriculum teaching syntactic embedding upon reading comprehension of fourth graders (Doctoral dissertation, The Florida State University, 1971). *Dissertation Abstracts International,* 1972, *32,* 4850A-4851A. (University Microfilms No. 72-10,049).

Stotsky, S. L. Sentence-combining as a curricular activity: Its effect on written language development and reading comprehension. *Research in the Teaching of English,* 1975, *9,* 30-68.

The Alberta Journal of Educational Research *Vol. XXIV, No. 2, June, 1978*

GEORGE S. C. CHEONG
Mount Saint Vincent University
and
EDGAR P. WADDEN
Lake View School

The Relationship Between Teachers' Experimental and Dogmatic Attitudes, and their Pupils' Self-Concept

The purpose of this study was to examine teachers' effect on fourth-, fifth-, and sixth-grade pupils' self-concept when they were taught by two extreme, significantly different groups of teachers: one group as the most experimental or the least dogmatic, and the other as the least experimental or the most dogmatic. The result of the study was that pupils who were taught by the most experimental group of teachers had significantly higher self-concept than pupils who were taught by the least experimental group of teachers. Implications were drawn from the study. (Dr. Cheong is in the Education Department at Mount Saint Vincent University, Halifax, and Mr. Wadden is Principal of Lake View School, Halifax County.)

In teacher effectiveness research, Mitzel (1960) suggested that three criteria be used: (a) presage variables (teacher personality, knowledge, status characteristics); (b) process variables (teacher behaviours, pupil behaviours, and teacher-pupil interactions); and (c) product variables (primarily measures of pupil change). Following Mitzel's theoretical model, McDonald and Elias (1976) investigated teacher effectiveness, and by means of stepwise regressions and path analyses, they found significant relations in teacher effectiveness and pupils' achievement in reading and mathematics. Also using Mitzel's model as the basis, Cantrell, Stenner, and Katzenmeyer (1977) combined all three criteria simultaneously, and their finding was that first-grade teachers characterized by high knowledge of behavioural principles were more verbally positive with their classes and produced significantly higher residual achievement gain results for low and middle IQ pupils than did first-grade teachers characterized by low knowledge of behavioural principles.

Rodin and Rodin (1972) researched teacher effectiveness and found that students rated most highly those faculty members from whom they learned the least. Cunningham's (1975) concern was the compatibility between teachers' teaching styles and learners' learning styles by using pupil achievement as the criterion; his findings were that certain types of teachers were significantly more effective with one type of student than with another, and that all student types seemed to have at least one type of teacher with whom they achieved significantly greater success. Apart from finding linear relationships between teacher behaviour and pupil performance, Clark, Snow, and Shavelson (1976), and Soar and Soar (1977) found in a few cases curvilinear relationships between teacher behaviour and pupil performance.

Costin, Greenough, and Menges (1971) reviewed numerous studies on the relationship between students' achievement and teacher behaviour, but the results were highly varied. Some investigations found positive correlations, others reported negative correlations, and still others found no significant relationships. A recent study by Armento (1977) is a typical example of generalisations summarized by Costin and his associates.

Using Protocol Modules dealing with sensitising teachers' awareness of pupil self-concept as the treatment, Borg (1977) found no significant improvement in pupil self-concept. However, this study, basically in agreement with Mitzel's (1960) theoretical framework, was concerned with the impact made by certain types of teacher personality attributes upon the acquisition of self-concept by pupils.

The kind of teacher personality attributes involved here is the experimental attitude expounded by John Dewey and later schematized by Brown (1968). Two implicit assumptions of the study which should be pointed out are that experimental attitude is a desirable trait for teachers to possess, and that teachers are the central figure in a formal teaching-learning situation.

The purpose of this study was to investigate the impact of the most experimental or the least dogmatic teachers and that of the least experimental or the most dogmatic teachers upon their pupils' acquisition of self-concept after an exposure of treatment for one semester. Therefore the hypothesis of the study was that grade four, five, and six pupils taught by teachers more experimental in attitude achieve significantly higher scores in self-concept than pupils taught by teachers less experimental in attitude.

To measure pupils' self-concept, Piers-Harris Children's Self-Concept Scale was used, which consists of 80 items. The authors (1964) reported split-half reliability coefficients ranging from .87 to .90, and KR 21 coefficients ranging from .78 to .93 for this measure. This instrument was developed to measure broad, general self-concept dealing with physical, social, academic, and emotional dimensions. Self-concept is derived from the child's own perceptions of himself (herself), the child's perceptions of how his (her) peers perceived him (her), the child's perceptions of how his (her) family perceived him (her), and the child's perceptions of how his (her) teacher perceived him (her).

Method

A city, a large town, a small town, and a rural area were involved in the study, giving a reasonably good indication of the cross-section for the

Province of Nova Scotia. However, it should not be misconstrued to mean that the sample was random and representative of population of the entire province.

By flipping a coin, Form A rather than Form B of Personal Beliefs Inventory and of Teacher Practices Inventory of the Experimentalism Scale, consisting of forty items each (Brown, 1968) was chosen for the study. And on the basis of Rokeach's (1956, 1960) results of an item analysis, the best forty items of the Dogmatism Scale were used in the study. The equivalent-forms and test retest reliability coefficients of Personal Beliefs Inventory of the Experimentalism (X) Scale range from .55 to .78, whereas those of Teacher Practices Inventory of the X Scale range from .56 to .94. The Dogmatism (D) Scale possesses internal consistency and test retest reliability coefficients ranging from .58 to .93. Therefore the reliability coefficients for both instruments were from moderate to high.

Before the 1972-73 academic year began, the Experimentalism Scale and the Dogmatism Scale were administered by the investigators to all grade four, five, and six teachers, except specialists in such subjects as Music, Art, French, and Physical Education, in their respective locations in early September of 1972. In total, 168 teachers participated.

One point that needs to be made explicit is that it was the experimental design of the study to exclude all specialist teachers because it was felt that they did not spend sufficient hours with pupils and since the study was designed to measure regular classroom teachers' impact on pupils. Schools which implemented subject teaching (one teacher teaching only one subject to upper elementary grades) were also excluded, again because it was felt that they did not spend sufficient hours with pupils. Teachers who had taught the same classes the previous year were excluded, since the study was designed to measure teachers' effect upon pupils for a specific period of time.

As soon as teachers' scores on the X Scale and the D Scale were available, eleven teachers with the highest scores on the X Scale and the lowest scores on the D Scale were chosen to constitute the *high* group of the study, and eleven teachers with the lowest scores on the X Scale and the highest scores on the D Scale were selected to constitute the *low* group. The reasons for truncating eleven pairs of teachers from the score continuum were: (a) the fact that there would be more than 500 pupils in the study, an adequate number for statistical purposes, and (b) the study was not funded by any agency. A t test (Guildford & Fruchter, 1973) was then performed; these two groups of teachers were significantly different on the X Scale ($t = 7.80$, df = 20, $p < .01$) and the D Scale ($t = 13.96$, df = 20, $p < .01$).

An informal examination was made of the twenty-two teachers chosen for the study with respect to their sex, age, and the school systems where they taught. It revealed that there was representation of male and female teachers, old and young teachers, and teachers teaching in the city system as well as those teaching in rural schools.

In mid-September of 1972, the Piers-Harris Children's Self-Concept Scale, "The Way I feel about Myself" (80 items) was administered by the investigators to twenty-two classes as the pre-test. In mid-February of 1973, the same scale was given to the same twenty-two classes ($N = 544$, with complete data) as the post-test.

Results

Data were analyzed by means of analysis of covariance (Nie, Hull, Jenkins, Steinbrenner & Bent, 1975) with the pupils' pretest scores on the Piers-Harris Children's Self-Concept Scale as the covariate. The results are presented in Table 1.

TABLE 1
ANALYSIS OF COVARIANCE ON FOURTH-SIXTH GRADE PUPILS' PRE- AND POSTTEST SCORES ON THE PIERS-HARRIS CHILDREN'S SCALE BETWEEN THE HIGH AND LOW GROUPS

Source of Variation	Sum of Squares	df	Mean Square	F
Covariates				
Pretest Scores	586.766	1	586.766	3.974*
Main Effects	4433.508	1	4433.508	30.028**
Residual	160194.941	1085	147.645	

* $p < .05$
** $pp < .001$

The *F* ratio was significant at .001 level. This finding indicates that pupils taught by the most experimental or the least dogmatic teachers had a significantly higher self-concept at the end of one semester than pupils taught by the least experimental or the most dogmatic teachers. Therefore the hypothesis was accepted.

Brown (1968) validated the X Scale against the criterion measure D Scale. As mentioned earlier, the X Scale comprises Personal Beliefs Inventory-Forms A and B, and Teacher Practices Inventory — Forms A and B. The concurrent validity coefficients ranged from .26 to .36. It is true that these validity coefficients, in spite of their being significant beyond .01 level, are quite moderate in magnitude. But it is generally recognized that substantial validity coefficients are difficult to obtain.

A second point related to the use of the pretest scores as the covariates, which requires some clarification, is that the pretest scores were inserted into the design to remove extraneous variation from the posttest scores, thereby increasing measurement precision. The finding indicated that a significant amount of variance in the posttest scores is attributable to the pretest scores. However, the computer program was designed in such a way that the analysis of variance was performed on the "corrected" scores, that is, "corrected" posttest scores which were our primary concern in terms of measuring the impact made by teachers in the *high* group and those in the *low* group upon their respective pupils' self-concept over a period of one semester.

Discussion

It might well be argued that children's self-concept is positively correlated with their school achievement. On the basis of the findings of this study, it can be safely concluded that the types of teachers, specifically in terms of their personality or to use Mitzel's scheme (1960), the relation between the

presage criteria (teachers' cognitive preparation and affective traits), the process criteria (the instructional process), and the product criteria (pupils' acquisition of cognitive skills and affective attitudes), had significant impact upon pupils' self-concept which is positively correlated with their achievement in school. It appears that it would be sensible to pay attention to the affective aspect in teacher preparation. Nevertheless, it is realized that there is greater difficulty in measuring validly and reliably the affective traits than the cognitive skills. Perhaps this is why, in teacher education heretofore, close attention seems to have been given to the cognitive aspects such as GPA and/or scores on the National Teacher Examination. It is now apparent that educators can no longer afford to ignore or to pay lip service to the affective dimension of teacher education. Since children's self-concept is a crucial determinant for their achievement in school, and the types of teacher personality do make a difference, as much attention as possible should be given to the affective dimension of teacher education programs as well as of teacher candidates.

References

Armento, B. J. Teacher behaviours related to student achievement on a social science concept test. *Journal of Teacher Education,* 1977, *28,* 46-52.

Borg, W. R. Changing teacher and pupil performance with Protocols. *Journal of Experimental Education,* 1977, *45,* 9-18.

Brown, B. B. *The experimental mind in education.* New York: Harper and Row, 1968.

Cantrell, R. P., Stenner, A. J., & Katzenmeyer, W. G. Teacher knowledge, attitudes, and classroom teaching correlates of student achievement. *Journal of Educational Psychology,* 1977, *69,* 172-179.

Clark, C. M., Snow, R. E., & Shavelson, R. J. Three experiments on learning to teach. *Journal of Teacher Education,* 1976, *27,* 174-180.

Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching and reliability, validity, and usefulness. *Review of Educational Research,* 1971, *41,* 511-535.

Cunningham, W. G. The impact of student-teacher pairings on teacher effectiveness. *American Educational Research Journal,* 1975, *12,* 169-189.

Guilford, J. P., & Fouchter, B. *Fundamental statistics in psychology and education* (5th ed.). New York: McGraw-Hill, 1973.

McDonald, F. J., & Elias, P. A report on the results of phase II of the beginning teacher evaluation study: An overview. *Journal of Teacher Education,* 1976, *27,* 315-319.

Mitzel, H. E. Teacher effectiveness. In C. W. Harris (Ed.), *Encyclopedia of educational research* (3rd ed.). New York: Macmillan, 1960.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. *Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill, 1975.

Piers, E. V., & Harris, D. B. Age and other correlates of self-concept in children. *Journal of Educational Psychology,* 1964, *55,* 91-95.

Rodin, M., & Rodin, B. Student evaluations of teachers. *Science,* 1972, *177,* 1164-1166.

Rokeach, M. Political an religious dogmatism: An alternative to the authoritarian personality. *Psychological Monograph,* 1956, *70* (18) (Whole No. 425).

Rokeach, M. *The open and closed mind.* New York: Basic Books, Inc., 1960.

Soar, R. S., & Soar, R. M. An attempt to identify measures of teacher effectiveness from four studies. *Journal of Teacher Education,* 1977, *27,* 261-267.

BOOK REVIEWS

NEW CAREER OPTIONS FOR WOMEN: A COUNSELOR'S SOURCEBOOK. *By Helen S. Farmer and Thomas E. Backer.* New York: Human Sciences Press, 1977, 349 pp., $16.95.

Educationalists have been aware, for some time now, of the way in which unquestioned cultural assumptions about sex-linked aptitudes and abilities have damaging effects upon the career choices made by, and the options made available to women as opposed to men. The bias that results from stereotypic assumptions affects both sexes, but in both personal terms and in terms of the issue of option and choice, greater limitations are placed upon women than upon men. A particularly disturbing observation is that those whose special task it is to advise people about their future labour market participation are often the worst offenders in the preservation and dissemination of outdated information and stereotypic, but unproven, assumptions about their clients' abilities based on their sex and thus, the labour market "niche" to which they might best be suited. Consequently, any material that might help change the status quo in both counselor attitudes and beliefs, and counseling techniques, is to be welcomed.

Farmer and Backer have produced a comprehensive review of women's position in the work world of the United States, but more importantly have produced a book that attempts to chart the "how to . . ." process in terms of the often special needs of women in their relation to the work world, and in terms of the counselor's own approach to the issues. The underlying assumption made by the authors is that all *persons* can achieve greater self-determination through the realization of the attitudes, behaviours and the social environment effects that may limit "optimal development and freedom to choose." In the pursuit of such an end the counseling process takes on a critical dimension irrespective of sex, but at the present time, most critically in terms of the female sex.

The authors initially undertake a review of the current position of women in the labour force in the United States, with some international comparisons and an examination and contradiction of some of the myths regarding "women's work." The presentation, in an admirably succinct manner, of current findings from sociological and psychological research is coordinated with references to inventory sources and testing programmes. This information comprises the first half of the book.

A major problem, however, is apparent in the presentation of the information mentioned above. Observing the concern of the authors, and of some practitioners in the field regarding existing counselor attitudes and bias, the presentation of information is remote for any discussion of the ways in which counselors might alter their attitudes and procedures. Such a criticism takes into account the problems any author confronts in organizing the material and information to be covered. In a book designed as a pragmatic aid to counselors, by-passing the opportunity to indicate the significance of the theories and facts presented for the day-to-day realities of the counseling situation and the concomitant problems faced by the client is a grave oversight. For example, there is a short section on child care, and some scattered references elsewhere to the issue of the working mother, but in the section devoted to the question of child care the authors conclude with the observation that, at present, "the available resources meet only 18 per cent of the need." Obviously any advice given to the working mother has to be given with this reality in mind, but the issue is not addressed by the authors in their subsequent discussion of the special needs of the working mother. For this reason, duplicated with regard to other particular concerns, needs and problems of the working women, it is probable that the book will be used by the majority of counselors more as a reference guide for programmes and local, professional, or union interest contacts, than for its initial presentation of the problems and needs of working women. Alternatively, others who relate theory and practice are likely to already comprise the "converted." A great pity, if such is the case, particularly in terms of the aforementioned concern with the documented prevalence of counselor bias with regard to "women's work."

In the presentation of the resources and agencies available to the counselor and client, Farmer and Backer have produced a well-organized catalogue for the U.S. In the sections dealing with specific counseling techniques and resources an array of interesting and, in some cases, innovative, ideas is outlined. But the discussion is guided by the assumption that the counselor has almost unlimited access to funds, and awesome power to effect change in a multitude of institutional settings. Such optimism on the part of the authors is both understandable and excusable given the assumptions underlying their efforts, but this does not alter the fact that such optimism is frequently misplaced in the reality of the counseling process. Perhaps in the pragmatic interest of the day-to-day counseling effort, some Nader-like investigation is required indicating the ways in which problems such as inadequate funding or even access to clients are frequent and might be overcome.

Assessing the book for Canadian education lists and counselors, it is estimated that about half of the book — that devoted to an assessment of current research findings and an examination of "special needs" — will be of direct interest. The remaining half of the book contains specific information of limited applicability in Canada, such as the text and discussion of some of the issues raised about the Equal Rights Amendment. In this context, one may conclude that a similar source book with more concern to the presentation of its material is long overdue for Canada.

R.A. Sydie
Department of Sociology
The University of Alberta

ADULT EDUCATION AS A FIELD OF STUDY AND PRACTICE. *By Duncan D. Campbell.* Vancouver: University of British Columbia Centre for Continuing Education. 1977, 230 pp., $5.00.

This book is a report of a study Mr. Campbell has recently completed on preparatory training for those involved in adult education. A review of such a study should examine its purposes, describe its procedures, and react to its outcomes.

This question emerges as one reads the report: Did the author start with a conceptualized program for the preparation of adult educators which he wished to test through the process of critical enquiry or was he searching for a theoretical framework within which to construct a model program? While the answer to this question is not clear and perhaps unimportant, one suspects that the study is intended to verify preconceived opinion.

The procedures followed in the study involved an extensive survey of the literature on adult education. The author has listed, under references cited, some hundred and seventy titles including books, journal articles, government documents and unpublished theses and papers. From these he has drawn a book-length series of direct quotations, paraphrasings and descriptions. He has organized this series around such themes as the product to be achieved through training, the processes and practices of training, and preferences which lead to the concluding recommendations of the report. The study is an excellent illustration of reflective enquiry. The challenge in this type of research is to achieve objectivity in selecting representative readings from the universe of existing literature on the topic and to analyze insightfully the opinions and practices expressed in those readings. On both counts Mr. Campbell has done a workmanlike job.

The author commences his study with a brief but adequate survey of adult education in Canada, an appropriate beachhead from which to attack the existing chaos of the field. He deals immediately with the diversity of definitions currently in use, each definition revealing a different perception of the present and future role of adult education. He does so by subsuming all of these definitions under the rubric of adult education thereby disposing of the need to examine the revolutionary implications of such descriptive terms as lifelong learning, nontraditional studies and community education. Otherwise he might have been forced to undertake three studies and write three books instead of one!

Mr. Campbell dismisses the European design in the training of adult educators which has been strongly influenced by the perception of adult education as a social instrument for improving the lot of powerless people. He chooses rather to view preparation as a "deepening of knowledge grounded in theory connected to and illumined by experience." This view leads inevitably to familiar patterns. His proposed training program is set within the contexts of psychology, sociology, philosophy and history with overtones of methodology, organization and administration.

Throughout the report the author stresses the need for research in the field of adult education. This he uses as his major argument for insisting that training programs for adult educators be located within universities. The conclusion is doubtless valid but hardly original. The same argument was used to justify the transfer of teacher education from the normal school to the university over thirty years ago. There were many more compelling reasons for this transfer historically than a pious recognition of the need for research.

While current writers in adult education are treading well worn paths,

these do not apparently lead to the same destination. A most interesting phenomenon appearing in the literature is the move to dissociate the preparation of teachers for adults from that of teachers for public and high school youth. Some universities in both North America and Europe are presenting what purports to be a new discipline, known as andragogy. This discipline stresses an instructional methodology designed for adult students, one which cannot become functional apparently in close association with those who are dedicated to the principles of pedagogy.

While Mr. Campbell refers only briefly to the discipline of andragogy he apparently agrees with its distinctions. His conclusion is that training for adult educators should not be set within regular faculties of education. To support this he describes practices followed in England where training in methodology is closely related to the subject of instruction. American practice is to locate training programs for adult educators within faculties of education. Nonetheless, some American writers cited by the author are opposed to this, insisting that it works to the disadvantage of adult educators-in-training. Coolie Verner, former head of the Centre for Continuing Education at the University of British Columbia, an independent unit, vigorously defends the practice in existence there. Mr. Campbell cites this defence in support of his recommendation.

It is difficult to establish rationally based evidence as justification for those positions or practices. The evidence presented in the report is loaded with value judgments. The methodology advocated by "andragogues" does not differ appreciably from that advocated by many teacher educators. For instance, the concept of teacher as a facilitator of learning rather than a director of instruction is not unknown to those who prepare teachers for the elementary and high school grades.

As an alternative, Mr. Campbell proposes that training programs for adults be placed with community development. This proposal has merit if one were to define adult education as a facet of community education, one phase only, albeit an important one, of community improvement. That adult education should become the handmaiden of community development does not appear among the assumptions of this report.

Some English writers identify the future role of adult education as encompassing recurrent periods of learning over the lifetime of an individual. If this should happen these writers forecast a paradigm shift in education. We now expect that formal education will terminate for most people well before the age of twenty-five. If each individual is guaranteed an entitlement of sixteen to eighteen years of formal schooling with the option of taking the post compulsory years according to personal preference we are presented with a completely different view of schooling arrangements. The conventional education-work-leisure-retirement sequence would be dramatically altered. Any move in this direction would further blur distinctions between preparatory and continuing education.

The study contains a compendium of knowledge about adult education. Many of Mr. Campbell's analytical interpretations of this wealth of information are exceedingly germane. He has made a useful contribution to adult education as a field of study.

T. C. Byrne, Visiting Professor
Department of Educational Administration
The University of Alberta

PEOPLE, CULTURE AND LEARNING. *By John W. Friesen.* Calgary: Detselig Enterprises Ltd., 1977, 255 pp., $12.25.

In *People, Culture and Learning,* the author, John W. Friesen, has set an ambitious goal of combining theory with practice within the area of intercultural education with the focus being on the " phenomenon of culture." If this book is judged strictly in terms of this goal, it will be seen as a less significant contribution than if it is more modestly viewed as an attempt to provide an awareness of and insight into the cultural distinctiveness of three ethnic groups in western Canada.

Friesen has broken the book into two parts. Part One, consisting of five chapters, is an attempt to provide a theoretical framework for an analysis of intercultural education. Part Two, including seven chapters, consists of case studies of the Plains Indians, Hutterites and Mennonites. A selected bibliography, organized according to topics, has been provided which will be useful to a student unfamiliar with the area of intercultural education.

That the book has been published by an unknown local enterprise may be due in part to the fact that a major proportion of the material is neither totally new nor unavailable. Of the twelve chapters, eight are essentially articles which have appeared between 1971 and 1976 in readily available journals.

As was indicated earlier, Friesen has failed mostly in dealing with the theoretical components of intercultural education. What he presents as the sociological base constitutes little more than an affective response to the importance of intercultural education. One may well agree with the humanistic conclusions about majority-minority relations that Friesen reaches, but the reasons for these conclusions are based more on a messianic commitment than a sociological analysis. Like so many other authors in this field, Friesen has relied on an orientation which is practical but, characteristically, atheoretical. Of the sociological and anthropological works cited, few represent contributions to the field, most citations being at least ten years old. Any reader approaching the book with a background in any of the social sciences will be less than satisfied with the theoretical conceptualization and the application of the "phenomenon of culture." Although Friesen has discussed dominant and variant value orientations in Chapter Five, he has made no attempt to provide an analytic discussion of the dynamics of culture. Furthermore, clarity in the use of concepts is lacking. For example, intercultural, multicultural and pluralistic are used interchangeably.

A final criticism of Part One of the volume is that it has failed to view majority-minority relations in a pluralist society within an arena of differentially distributed power. Chapter Two, concerning legal and theoretical considerations, which could have been the basis of a discussion of power, failed somewhat due to its descriptive and nonanalytic nature.

On a more positive note, Chapters Three, Four, and Five provide an interesting descriptive account of the development and growth of intercultural education and the role of values in intercultural education. However, these do not contribute to the theoretical component of the book as promised by the author and would have been more reasonably part of a general introduction to the volume.

Part Two is the book's saving grace. Here Friesen has discussed at length three religio-ethnic groups, the Plains Indians, the Hutterites and the Mennonites. Evaluating the treatment of the three groups according to the author's objectives, one must reach the conclusion that the most successful

case study is that of the Plains Indians. The strength of this analysis is in the fact that Friesen has discussed the Indians' characteristics extensively and, equally important, has looked at their needs, wants, and aspirations — a very important part of an understanding of cultural dynamics in a pluralistic society. It is this latter characteristic which is lacking in the treatment of the Mennonites and Hutterites. One learns who they are and how the majority groups in society view them but one gains little understanding of how they appraise their world in terms of the power dimension, with all its political and economic manifestations. Knowing this, one could more realistically approach the philosophical questions Friesen has raised in Part One of the volume. Taken together, the discussion of the three groups is informative and compassionate. The references for further reading in these areas are well chosen and will provide the reader with a readily available source of supplemental readings.

In conclusion, Friesen has presented a study of three religio-ethnic groups which has been done with concern and compassion. For the beginning teacher who must be sensitized to cultural differences between pupils who are likely to be members of different ethnic groups, *People, Culture and Learning* represents a readable account of some such differences. However, the serious student of intercultural education will find this volume lacking in that it presents no clear theoretical orientation which can be used in focussing on culture or on relations of groups within a pluralist society. Unfortunately, this book is likely, at best, to encourage toleration of differences between minority groups in the name of humanism but will provide little understanding of the dynamics of minority-majority group relations. Surely it is the latter that is needed in understanding the problems of and possibilities for Canadian society.

John R. Young
Associate Professor
Department of Educational Foundations
The University of Alberta

COGNITION, CURRICULUM, AND COMPREHENSION. *By John T. Guthrie (Ed.).* Newark, Delaware: International Reading Association, 1977, 300 pp., $8.50.

There are a number of difficulties in reviewing this book: (1) The articles contained in the book are constrained by the format of the conference where these papers were first presented. Therefore it might be more logical to review the conference than to review the resulting manuscript. (2) Rather than expressing unity of thought, the book contains ten independent papers, most of which are followed by a reaction paper. The final paper attempts to be a review of all the papers and their reactions.

Since evaluation only makes sense in terms of objectives, it is necessary to consider the objectives of this book — in actual fact, the consideration of the conference objectives. The conference known as "The Development of Reading Comprehension," was held in Newark in the summer of 1975. The Preface states that ". . . the conference was intended to explore basic research and the teaching of comprehension" (p. xiii). "Basic research" was not defined. An analysis of the papers, however, indicates that whereas some papers were based completely around a research study in which the author

was involved (Wilder, p. 257), other papers did not include any research by the author (Athey, p. 71). One may also question the purpose of the reactions of one author to the paper of another author. There does not appear to have been any guide-lines as to what constitutes a reaction. The reactions seem to range from a paraphrase of the paper with additional support, and alternate suggestions (Hansen, p. 65) to adulation (Carroll, p. 157) to an apparent attempt to insert one's own point of view (Perfetti, p. 279).

The Foreword is very difficult to understand and is rather pretentious and nondirective. "This volume represents a confluence of significant current trends in the scientific study of reading processes and practical aspects of curriculum design" (p. ix). "Highlighted in this book is the concern with the advanced competencies of reading and verbal literacy subsumed under the general name of comprehension" (p. x). ". . . the ineffectiveness of many curricular reforms . . . (is) based upon ignorance of the nature of the cognitive activities involved in acquiring knowledge from text and reasoning with written language . . . This point of view is the theme of this volume" (p. xi).

The individual papers range over a number of topics of interest to the authors and depending on the reader's background, will provide varying degrees of information. The recency of some of the research quoted may be questioned and recent parameters of discourse analysis (propositions, topic/comment, new/old information, staging) are lacking.

Carroll's (p. 1) is a general paper and he advises assessing independently the influence of reading, language, and cognition by means of a w - o - c scale, where w - o - c stands for written - oral - cognitive. However, he does not adequately distinguish these three components except that the last two are involved in the first. Perfetti (p. 20) does not present a conceptual definition of reading and at one point (p. 24) equates reading with phonological processing. His identification of Levels I, II, III do help the reader focus on certain components to be considered when children are reading. The choice of words in this article makes it difficult to read at times. For example, "Since time is continuous are we dealing with degrees of conscious effort or with some fairly sharp distinction between automatic and effortful" (p. 24) and "On the other hand it is adaptive to forget the words after a certain point" (p. 33). Hansen (p. 65) provides a good review to the Frase (p. 42) paper which states "that learning outcomes are strongly related to the properties of the goals given to readers and their own focus of attention" (p. 65). The Athey (p. 71) paper attempts to analyse the relation of syntax and semantics to reading. The majority of articles are devoted to syntactic studies, most of which are older and well-known. In reviewing these studies, she does not make an effort to distinguish between production and reception, oral and written, but tends to describe the studies in chronological order. Although she makes the statement that "syntactic development does not consist of a series of structures which evolve in a programmed sequence unrelated to the development of meanings" (p. 78), she does not criticize syntactic research from this point of view. And whereas later in the paper she talks about semantic constraints, she does not apply this to her discussion of the Chomsky study. Whereas children may have difficulty in understanding such structures as "The doll is easy to see," they may have no difficulty with "The chair is easy to see." Hood (p. 101) presents a good paper with much information, some based on her own research. However, she does not attempt to fit language development into a comprehension framework. Beck (p. 113) provides a good evaluation of the Sullivan Reading Series and demonstrates how some of its weaknesses may be overcome by examining and specifying

basic concepts of the reading process and curriculum design. The construction of the New Primary Grade Reading System (NRS) would have had greater transferability for other designers of curriculum had she been more explicit on the relationship of the components of the program to the sequence of steps indicated within the reading process model. Hansen and Lovitt (p. 160) discuss applied behaviour analysis methodology as a new approach for conducting research in reading. Some components (e.g., direct measurement) have long been a part of criterion reference testing. Unlike Beck, the authors of this paper do not have any theory of reading to guide their research interpretations. Rate appears to be a major criterion of successful reading. The results of some of their studies (regardless of the methodology) appear to be "givens." For example, "The ability to comprehend is related to grade level" (p. 167) and "When difficulty of the material increases the pupil's ability to read and comprehend gradually worsens" (p. 171). Otto (p. 193) argues for a skills approach in reading instruction. Whereas he makes such statements as "Once we concern ourselves with sensibly chosen skills" (p. 196) and " . . . sensibly selected skills" (p. 198), he does not indicate the research on which these skills were chosen and arranged hierarchically. Stauffer (p. 242) sets forth his views on reading as a cognitive act but these are based on his 1969 book, *Directing Reading Maturity as a Cognitive Process*. Wilder (p. 257) presents an information paper on the United States Office of Education's commissioning Educational Testing Service in 1971, "to plan and conduct a 'descriptive and analytic study of compensatory reading programs in U.S. public schools'" (p. 257). Conclusions must be considered in the light of the limitations of research design and these are adequately raised by Perfetti in his review of this paper. However, either the word "exemplary" is ambiguous or there is a contradiction in the conclusions, for Wilder states "The five schools were chosen because their reading programs were 'exemplary'", but "In short, nowhere among the five schools was there apparent to us a definable concentrated set of principles or procedures related to instruction in comprehension" (p. 277).

The final paper by Guthrie which refers to some of the ideas presented in the volume and to other existing research and draws implications for instruction, would have provided a much better conclusion for the book had it synthesized the key ideas presented by the text contributors, and focussed on the broader issue which is raised in the Preface — the interrelation between research and curriculum design (as opposed to instruction).

No rationale is provide for the sequence of papers which may be a factor in comprehensibility. One rationale might be to begin with language development and then focus on reading as language and develop a concept of reading comprehension. This might be followed by research studies on particular aspects of reading comprehension, to be followed by the implications of this knowledge for curriculum design. With this rationale the following sequence is recommended.

(1) Hood (p. 101) on language development.

(2) Carroll (p. 1) and Stauffer (p. 242) on defining comprehension.

(3) Athey (p. 71) on linguistic parameters of reading.

(4) Frase (p. 42), Hansen and Lovitt (p. 160) and Perfetti (p. 20) on their research in reading. Perfetti is last because he moves into decoding (pronunciation).

(5) Beck (p. 113), Otto (p. 193), and Wilder (p. 257) on curriculum design. Beck follows Perfetti because the reading curriculum she describes originated as a code emphasis reading program.

Perhaps the greatest difficulty in this book serving its purpose results from the lack of clarity between the conference (oral language) where the papers were first presented, and the text (written language) where the ideas are more widely disseminated. With regard to purposes of conferences, Carroll states " . . . to get participants to rethink their objectives and methods; to get them to learn new terminology and concepts; review the latest research findings . . . and go away refreshed with the thought that, on arrival back home, they can lick all the problems that previously seemed insoluble" (p. 159). Guthrie makes the point that "Communication between researchers and practitioners may be enhanced by posing three questions to each group: 1) Do you have messages for your counterparts in reading? 2) Are these messages consistent? 3) Are the messages important, that is, worthy as a basis for change?" (p. 283). Had Carroll's guidelines been given to the conference participants and Guthrie's to the participants as manuscript writers, the book would have fulfilled a more definite purpose.

William T. Fagan
Department of Elementary Education
The University of Alberta

List of Books Received

Bandura, A. *Social learning theory.* Englewood Cliffs, NJ: Prentice-Hall, Inc., 1977, 247 pp., $9.50 (hardcover), $5.95 (paper).

Bruno, J. E. *Educational policy analysis.* New York: Crane, Russak & Company, Inc., 1976, 275 pp., $21.50.

Carlton, R. A., Colley, L. A., & MacKinnon, N. J. *Education, change and society: A sociology of Canadian education.* Agincourt, Ont.: Gage Publishing, 1977, 530 pp., $12.95 (hardcover), $8.95 (paper).

Cohen, L. *Educational research in classrooms and schools.* New York: Harper & Row Publishers Inc. 1976, 426 pp., $13.80.

Edmonds, E., & Bessai, F. *Small rural schools on Prince Edward Island.* Charlottetown: Square Deal Publications, 1977, 122 pp., $4.95 (paper). ISBN 0-920078-10-9.

Farmer, G. M., Graham, E. J., & Jenkins, L. M. *Professional applications in typewriting.* Agincourt, Ont.: Gage Educational Publishing Ltd., 1977, 239 pp., $8.95 (hardcover). ISBN 0-7715-0886-7.

Fuller, R. *In search of the IQ correlation.* Stony Brook, NY: Ball-Stick-Bird Publications, Inc., 1977, 184 pp., $10.50.

Glenn, J. A. (Ed.). *Teaching primary mathematics.* New York: Harper & Row, 1977, 110 pp., $4.60 (paper).

Kidd, J. R., & Selman, G. R. (Eds.). *Coming of age: Canadian adult education in the 1960's.* Toronto: Canadian Association for Adult Education, 1977, 410 pp., $13.95 (hardcover), $8.95 (paper). ISBN 0-920056-0-24.

Kingston, A. J. (Ed.). *Toward a psychology of language and reading.* Athens, GA: The University of Georgia Press, 1977, 219 pp., $12.00.

Leitch, S. M. *A child learns to speak: A guide for parents and teachers of preschool children.* Springfield, IL: Charles C. Thomas, 1977, 91 pp., $8.50 (hardcover), $5.95 (paper).

O'Bryan, K. G., Reitz, J. G., Kuplowska, O. M. *Non-official languages: A study in Canadian multiculturalism.* Ottawa: Secretary of State, 1977, 274 pp., $4.75 (Canada), $5.70 (other countries).

Perrott, E. *Microteaching in higher education: Research, development, & practice* (SRHE Monograph 31). Guildford, England: Society for Research into Higher Education, 1977, 80 pp. ISBN 0-900868-58-9.

Prout, P. F. *Community schools in Canada.* Toronto: Canadian Education Association, 1977, 36 pp., $2.00 (paper).

Robinson, H. A. (Ed.). *Reading and writing instruction in the United States: Historical trends.* Newark, Delaware: International Reading Association, 1977, 81 pp., $4.00. ISBN 0-87207-855-8.

Winter, S. A., & Cox, E. W. *Competency-based instruction for exceptional children.* Springfield, IL: Charles C. Thomas, 1976, 143 pp., $12.50.

Zeitlin. *Kindergarten screening: Early identification of potential high-risk learners.* Springfield, IL: Charles Thomas, 1976, $14.95.

New Zealand Journal of
Educational Studies

Published twice-yearly by the New Zealand Council for Educational Research

The New Zealand Journal of Educational Studies provides an avenue of publication for essays, research findings and critical comments in any of the major fields of educational study in New Zealand and the Pacific area.

The current subscription rate is US $8.00
Business correspondence and subscriptions should be addressed to the Executive Officer, New Zealand Council for Educational Research, Box 3237, Wellington, New Zealand.

New Zealand Journal of Educational Studies

Some recent articles:

The Response of New Zealand State Secondary Schools to Increasing Maturity in their Pupils
Elizabeth M. Campbell

A Socio-Economic Index for the Female Labour Force in New Zealand
J.C. Irving and W.B. Elley

The Education of 'The Plain People': The Amish in America
W.J.D. Minogue

Classroom Architecture and Classroom Behaviours: ALook at the Change to Open-plan Schools in New Zealand
W.M. Gill

Natural Directions: The New Zealand Movement for Sexual Differentiation in Education During the Early Twentieth Century
Margaret Tennant

PRINTED BY
THE UNIVERSITY OF ALBERTA
PRINTING DEPARTMENT